Upper bounds of eavesdropper's performances in finite-length code with decoy method

Masahito Hayashi[1, *]

[1]ERATO-SORST Quantum Computation and Information Project, JST
5-28-3, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

Security formulas of quantum key distribution (QKD) with imperfect resources are obtained for finite-length code when the decoy method is applied. This analysis is useful for guaranteeing the security of implemented QKD systems. Our formulas take into account the effect of the vacuum state and dark counts in the detector. We compare the asymptotic key generation rate in presence of dark counts with that without.

PACS numbers: 03.67.Dd,03.67.Hk,03.67.-a,05.30.Jp

I. INTRODUCTION

The BB84 protocol proposed by Bennett and Brassard[1] in 1984 attracts attention as an alternative to modern cryptography based on complexity theory. Many efforts are devoted to searching for implementations of quantum communication channels for this purpose. The security of the original BB84 protocol can be trivially proved only when the quantum communication channel is noiseless. Since there is noise in any implemented quantum channel, it is needed to prove the security with the noisy channel, which has been proved by Mayers[2]. After his proof, many different proofs were reported. However, any implemented quantum channel, besides loss, also suffers imperfections in generating a single photon. That is, the sent pulse is given as a mixture of the vacuum state, the single-photon state, and the multi-photon state, and it is impossible for the sender (Alice) and the receiver (Bob) to identify the number of photon. In order to guarantee the security in such a case, the decoy method has been proposed[3, 4, 5, 6], in which different kinds of pulses are transmitted. However, these preceding researches did not provide security with the finite-length code, which is a basic requirement in practical settings. That is, there is no established method to evaluate quantitively the security of an implementable quantum key distribution (QKD) system.

On the other hand, modern cryptographic methods are required to evaluate its security quantitively. Hence, for the practical use of QKD, it is needed a theoretical analysis in order to present quantitive criteria for security and to establish the method to guarantee this criteria for the implemented QKD system. If nothing in this direction is done, QKD systems cannot be developed for practical use.

In a usual QKD protocol, the final key is generated via classical error correction and privacy amplification after the initial key (raw key) is generated by the quantum communication. In the classical error correction part, it is sufficient to choose our classical error correction code based on the detected error rate. Privacy amplification, on the other hands, sacrifices several keys in order to guarantee the security against the eavesdropper. The upper bound of eavesdropper(Eve)'s information for the final key is closely related to the amount of sacrifice bits.

Since Eve's information for the final key is the measure of the possibility of eavesdropping, its quantitive evaluation is required. In order to decrease Eve's information sufficiently, we need a sufficient amount of sacrifice bits, which is given by the product between the length of our code and the rate of sacrifice bits. A larger size of our code requires larger complexity of the privacy amplification, and a larger rate of sacrifice bits decreases the generation rate of the final key. Hence, it is required to derive the formula to calculate the upper bound of Eve's information for the final key for a given length of the code and a given rate of sacrifice bits, under the realizable quantum communication channel.

Our problems can be divided into three categories: The first is the evaluation of Eve's information for the given length of our code and the given rate of sacrifice bits. Since any implemented QKD system has a finite-length code, any asymptotic security theory cannot guarantee the security of an implemented QKD system. The second is the security analysis for imperfect resource (e.g., phase-randomized coherent light) that consists of mixtures of the vacuum state, single-photon state, and multi-photon state. Many practical QKD systems are equipped not with single-photon but with weak phase-randomized coherent signals. These systems require a security analysis with an imperfect resource. Further, even if a QKD system is approximately equipped with single-photon signals, it nonetheless needs a security analysis for an imperfect resource because only a perfect single-photon resource allows the security analysis for the single-photon case. The third is the identification of the relative ratio among the vacuum state, the single-photon state, and the multi-photon state in the detected pulses. Many implemented quantum communication channels are so lossy that Alice and Bob cannot identify this ratio in the detected pulses even though they know this ratio in the transmitted pulses. Thus, they need a method to estimate this ratio. Each of these three problems has been solved only separately, however, an implemented QKD system requires a unified solution for these three prob-

*Electronic address: masahito@qci.jst.go.jp

lems, which cannot be obtained by a simple combination of separate solutions.

Concerning the first problem, many papers treated only the asymptotic key generation (AKG) rate. Only the papers, Mayers[2], Inamori-Lütkenhaus-Mayers(ILM)[7], S.Watanabe-R.Matsumoto-Uyematsu(WMU)[8], Renner-Gisin-Kraus(RGK)[9], and Hayashi[10] discussed the security of the finite-length code with a low complexity protocol. In particular, only ILM[7] takes into account the second problem among them, and the other papers treat only the single photon case. Extending the method of Mayers[2], ILM [7] provided an evaluation of the security with imperfect resources for the finite-length code. Unfortunately, their formula for the security evaluation is so complicated that a simpler security bound is needed. They also obtained the AKG rate with imperfect resources. Extending the method of Shor-Preskill[11], Gottesman-Lo-Lütkenhaus-Preskill (GLLP)[12] also obtained this rate. In order to solve the third problem, Hwang[3] developed the decoy method, in which we estimate the ratio by changing the intensity of the transmitted phase-randomized coherent light randomly. After this breakthrough, applying the asymptotic formula by GLLP[12], Wang[6] and Lo et al.[4, 5] analyzed this method deeply, but did not treat the security of the finite-length code. Hence, there is not enough results to treat the security with the decoy method for the finite-length code.

Further, there is a possibility for an improvement of the AKG rate by ILM [7] and GLLP[12]. Taking into account the effect of the vacuum state, Lo[13] conjectured an improved AKG rate. Considering the effect of the dark counts in the detector, Boileau-Batuwantudawe-Laflamme (BBL)[14] conjectured a further improvement of the AKG rate conjectured by Lo[13]. They pointed out that the AKG rate with the forward error correction is different from that with the reverse error correction.

In this paper, in order to evaluate eavesdropper's performances, we focus on the average of Eve's information, the average of the maximum of trace norm between Eve's states corresponding to different final keys, and the probability that Eve can correctly detect the final key. We derive useful upper bounds of these quantities for the protocol given in section II in the finite-length code by use of mixing different imperfect resources. Based on this bound, we obtain an AKG rate. In particular, due to the consideration of the effect of the dark counts, our bound improves that by ILM [7], and it yields the AKG rate that coincides with the that conjectured by BBL[14]. We should mention here that our description for quantum communication channel is given as a TP-CP map on the two-mode bosonic system. Since our results can be applied to the general imperfect sources, it provides security with an approximate single-photon source. However, further statistical analysis is required for the numerical bound of Eve's information for implemented QKD system with the finite-length code. Such an analysis is presented in another paper [15]. Also, the analysis of the AKG rate in the case of phase-randomized coherent light will be presented in another paper [16].

The paper is organized as follows. In section II, as a modification of BB84 protocol, we present our protocol, in which we clarify the measuring data deciding the size of sacrifice bits in the privacy amplification. In section III, we derive upper bounds of the averages of Eve's information about the final key and of the trace norm of the maximum between Eve's states corresponding to different final keys under the protocol given in section II. In section IV, we characterize the AKG rate based on our bounds, and apply it to the case of mixture of the vacuum and the single-photon and the case of approximate single-photon. In Section V, the quantum communication channel is treated as a general TP-CP map on the two-mode bosonic system. It is proved that such a general case can be reduced to the case given in section III.

II. MODIFIED BB84 PROTOCOL WITH DECOY STATE

We consider BB84 protocol based on $+$ basis, $|\uparrow\rangle$, $|\downarrow\rangle$ and $\times$ basis, $|+\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$, $|-\rangle := \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle)$. If we realize this protocol by using photon (or bosonic particle), we have to generate single-photon in the two-mode system and transmit it without no loss. However, it is impossible to implement this protocol perfectly, any realized quantum communication system can send only an imperfect photon (or approximately single-photon). Hence, we have to treat bosonic system more carefully. Let us give its mathematical description. Two-mode bosonic system is described by $\mathcal{H} := \oplus_{n=0}^{\infty}\mathcal{H}_n$, where the n-photon system $\mathcal{H}_n$ is the Hilbert space spanned by $|0,n\rangle, |1,n-1\rangle, \ldots, |n-1,1\rangle$, and $|n,0\rangle$. For example, $|j,n-j\rangle$ is the state consisting of j photons with the state $|\uparrow\rangle$ and $n-j$ photons with the state $|\downarrow\rangle$. Also the vector $\sum_{j=0}^{n}\sqrt{\binom{n}{j}(\frac{1}{2})^n}|j,n-j\rangle$ $(\sum_{j=0}^{n}\sqrt{\binom{n}{j}(\frac{1}{2})^n}(-1)^{n-j}|j,n-j\rangle)$ corresponds to the state of n photons with the state $|+\rangle$ ($|-\rangle$). That is, the system $\mathcal{H}_n$ is equivalent with the n-th symmetric subspace of two dimensional system. We also denote the state of j photons with the state $|+\rangle$ and $n-j$ photons with the state $|-\rangle$ by $|j,n-j,\times\rangle$.

When we would generate the state $|\uparrow\rangle$ in the two-dimensional system with the coherent pulse, the generated state is described by the state $\sum_{n=0}^{\infty} e^{-\frac{|\alpha|^2}{2}}\frac{\alpha^n}{\sqrt{n!}}|n,0\rangle$ in the two-mode bosonic system. However, if we implement our system so that each phase factor θ of the complex amplitude $\alpha = \sqrt{\mu}e^{i\theta}$ is completely random, our state can be regarded as the mixed state $e^{-\mu}\sum_{n=0}^{\infty}\frac{\mu^n}{n!}|n,0\rangle\langle n,0|$, which depends only on the intensity μ. In the following, we consider a more general case, in which the pulse sent by the sender (Alice) are given by $\rho_{0,+}^{\nu} := \sum_{n=0}^{\infty}\nu(n)|n,0\rangle\langle n,0|$, $\rho_{1,+}^{\nu} := \sum_{n=0}^{\infty}\nu(n)|0,n\rangle\langle 0,0|$,

$\rho^{\nu}_{0,\times} := \sum_{n=0}^{\infty} \nu(n)|n,0,\times\rangle\langle n,0,\times|$, $\rho^{\nu}_{1,\times} := \sum_{n=0}^{\infty} \nu(n)|0,n,\times\rangle\langle 0,n,\times|$, where ν is an arbitrary distribution.

Since our communication channel is lossy, the receiver (Bob) cannot necessarily detect all of the sent pulses. If the breakdown of the detected pulses (the ratio among the vacuum state, the single-photon state, n-photon state, and so on) is known, we can guarantee the security of BB84 protocol based on the discussion on subsection III B. However, since the usual quantum communication channel is lossy, there is a possibility that Eve can control the loss depending on the number of the photons. Hence, it is impossible to identify the the loss of each number of the photons if Alice sends the pulse by using one distribution ν. One solution is the decoy method[3, 4, 6], in which Alice randomly chooses the distribution ν and estimates the loss and the error probabilities of each number state. It is effective to choose the vacuum pulse $|0\rangle\langle 0|$.

In the following, we describe our protocol. First, we fix the following; the size N of our code, the maximum number $\overline{N}$ and the minimum number $\underline{N}$ of final key size, the number N' of sent pulses, the k distributions $\nu_1,\ldots,\nu_k$ of the generated number of photons, and the distribution ν_{i_0}, whose pulse generates the raw keys. Since the vacuum pulse and two bases are available, Alice sends $2k+1$ kinds of pulses, where the 0-th kind of pulse means the vacuum pulse, the i-th kind of pulse means the pulse with the $\times$ basis generated by the distribution ν_i, and the $i+k$-th kind of pulse means the pulse with the $+$ basis generated by the distribution ν_i for $i=1,\ldots,k$. For this purpose, they fix the probabilities $\overline{p}_0,\ldots,\overline{p}_{2k}$, and Alice generates i-th kind of pulse with the probability $\overline{p}_i$ for $i=0,\ldots,2k$. The probabilities $\overline{p}_{i_0}$ and $\overline{p}_{k+i_0}$ should be larger because these generate the pulses producing the raw keys. In this paper, we use the bold style for describing the vector concerning the index i representing the kind of pulse, as $\overline{\mathbf{p}} = (\overline{p}_0,\ldots,\overline{p}_{2k})$.

Before the quantum communication, they check the probability p_D of dark counts in the detector, and the probability p_S of errors of the $\times$ basis occurred in the detector or the generator, which can be measured by the error probability when the quantum communication channel has no error. Similarly, they the probability $\tilde{p}_S$ of errors of the $+$ basis occurred in the detector or the generator.

1. Alice sends the N' pulses, where each pulse is chosen among $2k+1$ kinds of pulses. She denotes the number of the i-th kind of pulses by A_i. ($\sum_{i=0}^{2k+1} A_i = N'$)

2. After sending N' pulses, Alice announces the kind of the each pulse (the basis and the distribution ν_i) by using public channel.

3. Bob records the numbers $C_0,\ldots,C_{2k}$ of detected pulses and the numbers $E_1,\ldots,E_{2k}$ of detected pulses with the common basis for each kind $i=0,\ldots,2k$ of pulses. Bob announces the positions of pulses with the common basis and the above numbers by using public channel.

4. Alice chooses $E_{i_0}-N$ bits among i_0-th pulses with the common basis and $E_{i_0+k}-N$ bits among the i_0+k-th pulses with the common basis, and announces these positions and their bit by using public channel. Bob records the number of errors as H_{i_0}, H_{i_0+k} and announces them by using public channel. If $E_{i_0} \le N$ or $E_{i_0+k} \le N$, they stop their protocol and return to the first step.

5. Alice and Bob announce their bit of the remaining kinds $i \neq 0, i_0, i_0+k$ of pulses, and record the number of error by H_i.

6. Using these informations, they decide the rates $\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N})$ and $\eta(\frac{H_{i_0}}{E_{i_0}-N})$ of error correction and the sizes of sacrifice bits $m(\mathcal{D}_i,\mathcal{D}_e)$ and $\tilde{m}(\tilde{\mathcal{D}}_i,\mathcal{D}_e)$ in the privacy amplification for the remaining the i_0-th kind of pulses and and the i_0+k-th kind of pulses, respectively, where we abbreviate the initial data $(\mathbf{A},\boldsymbol{\nu},p_S,p_D)$ and $(\mathbf{A},\boldsymbol{\nu},\tilde{p}_S,p_D)$ and the experimental data $(\mathbf{C},\mathbf{E},\mathbf{H})$ to $\mathcal{D}_i$, $\tilde{\mathcal{D}}_i$, and $\mathcal{D}_e$. If $N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N}) - m(\mathcal{D}_i,\mathcal{D}_e) < \underline{N}$ or $N\eta(\frac{H_{i_0}}{E_{i_0}-N}) - \tilde{m}(\tilde{\mathcal{D}}_i,\mathcal{D}_e) < \underline{N}$, they stop their protocol and return to the first step. Further, if $\overline{N} < N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N}) - m(\mathcal{D}_i,\mathcal{D}_e)$, they replace $m(\mathcal{D}_i,\mathcal{D}_e)$ by $N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N}) - \overline{N}$. Similarly, if $\overline{N} < N\eta(\frac{H_{i_0}}{E_{i_0}-N}) - \tilde{m}(\tilde{\mathcal{D}}_i,\mathcal{D}_e)$, they replace $\tilde{m}(\tilde{\mathcal{D}}_i,\mathcal{D}_e)$. by $N\eta(\frac{H_{i_0}}{E_{i_0}-N}) - \overline{N}$.

7. They perform N bits error correction for $+$ basis, and generate $N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N})$ bits.

8. They perform privacy amplification for the $+$ basis, and generate $N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N}) - m(\mathcal{D}_i,\mathcal{D}_e)$ bits.

9. They perform N bits error correction for the $\times$ basis, and generate $N\eta(\frac{H_{i_0}}{E_{i_0}-N})$ bits.

10. They perform privacy amplification for the $\times$ basis, and generate $N\eta(\frac{H_{i_0}}{E_{i_0}-N}) - \tilde{m}(\tilde{\mathcal{D}}_i,\mathcal{D}_e)$ bits.

If Bob detects both events $|0\rangle$ and $|1\rangle$ in the measurement of the $+$ basis, he decides one event with the probability $\frac{1}{2}$. In the following, this measurement is described by the POVM $\{M_\emptyset, M_0, M_1\}$.

Error correction (7., 9.) In the step 7. and 9., Alice and Bob generate $l+m$ bits with negligible errors from N bits X and X' by using one of the following protocols: (For example, l and m are choosen as $l+m = N\eta(\frac{H_{i_0+k}}{E_{i_0+k}-N})$ and $m = m(\mathcal{D}_i,\mathcal{D}_e)$.)

Forward error correction They share $N \times (l+m)$ binary matrix M_e. Alice generates other $l+m$ bits random number Z, and sends $M_e Z + X$ to Bob. Bob applies the decoding of the code M_e to the bits $M_e Z + X - X'$ to extract Z, and obtain Z'.

Reverse error correction Bob generates other $l+m$ bits random number Z, and sends $M_e Z +' X$ to Alice. Alice applies the decoding of the code M_e to the bits $M_e Z + X' - X$ to extract Z, and obtain Z'.

As mentioned later, since this error correction corresponds to a part of the twirling operation, their channel can be regarded as a Pauli channel from Alice to Bob in the forward case (from Bob to Alice in the reverse case).

Privacy amplification (8., 10.) In the step 8. and 10., Alice and Bob generate l bits from $l+m$ bits Z by using the following protocol. First, they generate the same $l \times (l+m)$ binary matrix M_p with the following condition:

$$\mathrm{P}\{Z \in \operatorname{Im} M_p^T\} \le 2^{-m} \tag{1}$$

for any non-zero $l+m$ bit sequence Z. Next, they generate l bits $M_p Z$ from $l+m$ bits Z.

Hence, combining the above error correction and the above privacy amplification, Alice can be regarded to send information by the code $\operatorname{Im} M_e / M_e(\operatorname{Ker} M_p)$.

The preceding researches[2, 7, 8] analyze the security when the binary matrix M_p for privacy amplification is chosen completely randomly. If we choose the binary matrix M_p by the Toeplitz matrix[17, 18], we need less random number. This is because Toeplitz matrix requires only $l+m-1$ bits random number while completely random binary matrix M_p does $(l+m)l$ bits random number. An $l \times (l+m)$ binary matrix $(\mathbf{X}, I)$ is called Toeplitz matrix[17, 18] when its element $\mathbf{X} = (X_{i,j})$ is given by $l+m-1$ random variables $Y_1, \ldots, Y_{l+m-1}$ as

$$X_{i,j} := Y_{i+j-1}.$$

Theorem 1 *Toeplitz matrix satisfies the condition (1) for any element $Z \neq 0 \in \mathbf{F}_2^{l+m}$.*

For a proof, see Appendix A.

III. EVALUATION OF EVE'S INFORMATION CONCERNING FINAL KEY

A. Formulation of channel

In this section, we assume a simplified Eve's attack, and evaluate the security against Eve's attack. In Section V, we will treat the general case of Eve's attack, and prove that the general case can be reduced to the case of this section.

First, we assume that Eve can distinguish the four states $|n,0\rangle$, and $|0,n\rangle$ in $+$ basis and $|n,0,\times\rangle$,$|0,n,\times\rangle$ in $\times$ basis. Hence, the input system can be described by N-th tensor product system $\mathcal{H}^{\otimes N}$ of $\mathcal{H} := \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus (\oplus_{n\ge 2}\mathcal{H}_{n,+}) \oplus (\oplus_{n\ge 2}\mathcal{H}_{n,\times})$, where $\mathcal{H}_0$ is the one-dimensional space spanned by $|0,0\rangle$, $\mathcal{H}_1$ is the two-dimensional space spanned by $|0,1\rangle$ and $|1,0\rangle$, $\mathcal{H}_{n,+}$ is the two-dimensional space spanned by $|n,0\rangle$, and $|0,n\rangle$, and $\mathcal{H}_{n,\times}$ is the two-dimensional space spanned by $|n,0,\times\rangle$ and $|0,n,\times\rangle$. The output system is described by N-th tensor product space $\mathcal{H}^{\otimes N}$ of $\mathcal{H}_0 \oplus \mathcal{H}_1$.

Then, the quantum communication channel from Alice to Bob is given by

$$\bigoplus_{\vec{n}} \sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e}) (\otimes_{i=1}^N \mathcal{E}_{e_i|n_i})(\rho), \tag{2}$$

where $\mathcal{P}_{\vec{n}}(\vec{e})$ is the distribution of $\vec{e}$ when $\vec{n}$ is fixed. Such a channel is called Pauli channel. Here, $\vec{n}$ and $\vec{e}$ are given as follows: Each element n_i of $\vec{n} = (n_1, \ldots, n_N)$ is chosen among $0, 1, (2,+), (2,\times), \ldots$. Each element e_i of $\vec{e} = (e_1, \ldots, e_N)$ is chosen as v or s when n_i is 0. It is chosen among $v, (0,0), (0,1), (1,0), (1,1)$ when n_i is 1. Otherwise, it is chosen among $v, 0, 1$. When n_i is 0 or 1, the channel $\mathcal{E}_{e_i|n_i}$ is defined as

$$\mathcal{E}_{e_i|0}(\rho) := \begin{cases} \langle 0,0|\rho|0,0\rangle |0,0\rangle\langle 0,0| & \text{if } e_i = v \\ \langle 0,0|\rho|0,0\rangle \rho_{mix,1} & \text{if } e_i = s \end{cases}$$
$$\mathcal{E}_{e_i|1}(\rho) := \begin{cases} |0,0\rangle\langle 0,0| \operatorname{Tr} P_{\mathcal{H}_1}\rho P_{\mathcal{H}_1} & \text{if } e_i = v \\ \mathsf{W}^{e_i} P_{\mathcal{H}_1}\rho P_{\mathcal{H}_1} (\mathsf{W}^{e_i})^\dagger & \text{otherwise,} \end{cases}$$

where

$$\rho_{mix,1} := \frac{1}{2}(|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|),$$
$$\mathsf{W}^{(x,z)} := \mathsf{X}^x \mathsf{Z}^z,$$
$$\mathsf{X}|0,1\rangle = |1,0\rangle,\ \mathsf{X}|1,0\rangle = |0,1\rangle,$$
$$\mathsf{Z}|0,1\rangle = -|0,1\rangle,\ \mathsf{Z}|1,0\rangle = |1,0\rangle.$$

When n_i is not 0 or 1, the channel $\mathcal{E}_{e_i|n_i}$ is defined as

$$\mathcal{E}_{e_i|n_i}(\rho) := \begin{cases} |0,0\rangle\langle 0,0| \operatorname{Tr} \rho P_{\mathcal{H}_{n_i}} & \text{if } e_i = v \\ \langle 0,n_i|\rho|0,n_i\rangle|0,n_i\rangle\langle 0,n_i| + \langle 1,n_i|\rho|1,n_i\rangle|1,n_i\rangle\langle 1,n_i| & \text{if } e_i = 0 \\ \langle 0,n_i|\rho|0,n_i\rangle|1,n_i\rangle\langle 1,n_i| + \langle 1,n_i|\rho|1,n_i\rangle|0,n_i\rangle\langle 0,n_i| & \text{if } e_i = 1. \end{cases}$$

The raw key is generated from detected pulses, which belong to the system $\mathcal{H}_1$ on the Bob's side. Thus, we focus only on the pulse whose measurement value is not v. In the following, we consider the security from the final key distilled from raw keys of the + basis. Hence, the generated state can be restricted to $\mathcal{H}_0 \oplus \mathcal{H}_1 \oplus (\oplus_{n\geq 2}\mathcal{H}_{n,+})$. Thus, it can be assumed that our channel $\bigoplus_{\vec{n}} \sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}_{e_i|n_i})$ satisfies that each element e_i of $\vec{e} = (e_1,\ldots,e_N)$ is not v.

B. Security of known channel: no dark count case

Assume that the input state belongs to the subsystem $\mathcal{H}_{\vec{n}} := \mathcal{H}_{n_1}\otimes\cdots\otimes\mathcal{H}_{n_N}$ labeled by $\vec{n} = (n_1,\ldots,n_N)$. Now, we classify the N input subsystems into three parts:

0th part: $K^0(\vec{n}) := \#\{i|n_i = 0\}$.

1st part: $K^1(\vec{n}) := \#\{i|n_i = 1\}$.

2nd part: $K^2(\vec{n}) := \#\{i|n_i \geq 2\}$.

In the 0-th part, Eve can obtain no information. That is, Eve's information is equal to Eve's information when the Alice's information is sent by the + basis via the qubit channel:

$$\mathcal{E}'_{s|0}(\rho) := \frac{1}{2}(\mathsf{X}^0\rho(\mathsf{X}^0)^\dagger + \mathsf{X}^1\rho(\mathsf{X}^1)^\dagger).$$

In the 2nd part, Eve can obtain all of Alice's information by the following method: Eve receives two-photon state. She sends one qubit system to Bob, and keeps the other qubit. After the announcement of the basis, Eve measures her system with the correct basis. Thus, Eve's information is equal to Eve's information when the Alice's information is sent by the + basis via the phase-damping qubit channel (pinching channel):

$$\mathcal{E}'_{e_i|(n,+)}(\rho) := \frac{1}{2}(\mathsf{Z}^0\rho(\mathsf{Z}^0)^\dagger + \mathsf{Z}^1\rho(\mathsf{Z}^1)^\dagger)$$

for $n \geq 2$. This is because the channel is given by $\mathcal{E}'_{e_i|(n,+)}$ in the single photon case when Eve measures the system with the correct basis. Here, the presence or the absence of the error X is not so important for Eve's information. This is because the probabilities concerning the action Z is essential, as is discussed in Appendix C. Therefore, Eve's information concerning total N bits is equal to Eve's information when the N bits information $x_1,\ldots,x_N$ is sent by $|x_1,\ldots,x_N\rangle \in (\mathbb{C}^2)^{\otimes N}$ via the following qubit channel:

$$\sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}'_{e_i|n_i}). \tag{3}$$

There is a relation between the error probability in the $\times$ basis and the security.

Theorem 2 *Define $P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}$ as the error probability by an arbitrary decoding when an information sent by the code $(M_e(\operatorname{Ker} M_p))^\perp/(\operatorname{Im} M_e)^\perp$ with the $\times$ basis via the qubit channel $\sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}'_{e_i|n_i})(\rho)$, where $\vec{n} := (n_1,\ldots,n_N)$. When Alice sends l bits information with the code $\operatorname{Im} M_e/M_e(\operatorname{Ker} M_p)$ in the + basis via the same channel, the following relations hold.*

Define $\rho^E_{[Z]|M_p}$ as the final Eve's state when Alice's information is $[Z] = M_p Z$. Then, Eve's information $I^{\mathcal{P}_{\vec{n}}}_{E|M_p}$ is given as the quantum mutual information:

$$I^{\mathcal{P}_{\vec{n}}}_{E|M_p} := \frac{1}{2^l}\sum_{[Z]} D(\rho^E_{[Z]|M_p}\|\overline{\rho}^E_{M_p})$$
$$D(\rho\|\rho') := \operatorname{Tr}\rho(\log\rho - \log\rho')$$
$$\overline{\rho}^E_{M_p} := \sum_{[Z]} \frac{1}{2^l}\rho^E_{[Z]|M_p}.$$

The quantum mutual information $I^{\mathcal{P}_{\vec{n}}}_{E|M_p}$ satisfies that

$$I^{\mathcal{P}_{\vec{n}}}_{E|M_p} \leq \overline{h}(P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}) + l P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}, \tag{4}$$

where $\overline{h}(x)$ is defined as

$$\overline{h}(x) := \begin{cases} -x\log_2 x - (1-x)\log_2(1-x) & \text{if } 0 \leq x \leq 1/2 \\ 1 & \text{if } 1/2 < x \leq 1. \end{cases}$$

Hence, Eve's information per one bit is evaluated as

$$\frac{I^{\mathcal{P}_{\vec{n}}}_{E|M_p}}{l} \leq \frac{\overline{h}(P^{\mathcal{P}_{\vec{n}}}_{ph|M_p})}{l} + P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}. \tag{5}$$

We also obtain the following:

$$\min_{[Z]\neq[Z']} F(\rho^E_{[Z]|M_p}, \rho^E_{[Z']|M_p}) \geq 1 - 2P^{\mathcal{P}_{\vec{n}}}_{ph|M_p} \tag{6}$$
$$\max_{[Z]\neq[Z']} \|\rho^E_{[Z]|M_p} - \rho^E_{[Z']|M_p}\|_1 \leq 4P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}. \tag{7}$$
$$\min_{[Z]} F(\rho^E_{[Z]|M_p}, \overline{\rho}^E_{M_p}) \geq 1 - P^{\mathcal{P}_{\vec{n}}}_{ph|M_p} \tag{8}$$
$$\max_{[Z]} \|\rho^E_{[Z]|M_p} - \overline{\rho}^E_{M_p}\|_1 \leq 2P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}, \tag{9}$$

where $F(\rho,\rho') := \mathrm{Tr}\sqrt{\sqrt{\rho'}\rho\sqrt{\rho'}}$. Define $P^{\mathcal{P}_{\vec{n}}}_{succ|M_p}$ as the probability of successfully detecting the Alice's information $[Z]$. Then, the inequality

$$P^{\mathcal{P}_{\vec{n}}}_{succ|M_p} \le \left(\sqrt{P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}}\sqrt{1-2^{-l}} + \sqrt{1-P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}}\sqrt{2^{-l}}\right)^2 \tag{10}$$

holds.

Further, the concavity of $\overline{h}$ implies that

$$\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} I^{\mathcal{P}_{\vec{n}}}_{E|M_p} \le \overline{h}(\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}) + l\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p} \tag{11}$$

$$\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} \frac{I^{\mathcal{P}_{\vec{n}}}_{E|M_p}}{l} \le \frac{\overline{h}(\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p})}{l} + \mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}. \tag{12}$$

The concavity of left hand side of (10) holds concerning $P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}$. Thus,

$$\begin{aligned}&\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{succ|M_p}\\ \le& \left(\sqrt{\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}}\sqrt{1-2^{-l}} + \sqrt{1-\mathrm{E}^{\mathcal{P}_{\vec{n}}}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph|M_p}}\sqrt{2^{-l}}\right)^2.\end{aligned}$$

For a proof, see Appendix C. As shown in Section V, sending l bits information with the code $\mathrm{Im}\, M_e/M_e(\mathrm{Ker}\, M_p)$ is equivalent with the combination of sending random number and forward error correction by $\mathrm{Im}\, M_e$ and privacy amplification by $M_e(\mathrm{Ker}\, M_p)$.

Next, we focus on the average error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min|M_p}$ with the minimum length decoding when an information sent by the code $(M_e(\mathrm{Ker}\, M_p))^\perp/(\mathrm{Im}\, M_e)^\perp$ with the $\times$ basis via the qubit channel $\Lambda(\rho) = \sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}'_{e_i|n_i})(\rho)$. This value is described as

$$P^{\mathcal{P}_{\vec{n}}}_{ph,min|M_p} = \frac{1}{|(\mathrm{Im}\, M_e)^\perp|}\sum_{z\in(\mathrm{Im}\, M_e)^\perp}\sum_{z':(13)} \langle z'|\mathcal{P}_{\vec{n}}(|z\rangle\langle z|)|z'\rangle,$$

where we take the summand concerning z' satisfying the following condition (13):

$$\underset{z''\in(M_e(\mathrm{Ker}\, M_p))^\perp}{\mathrm{argmin}} \mathrm{dis}(z',z'') \in (\mathrm{Im}\, M_e)^\perp, \tag{13}$$

and $\mathrm{dis}(z',z'')$ is the Hamming distance between z' and z''. In order to analyze the error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min|M_p}$, we introduce the number $t(\vec{e},\vec{n})$:

$$t(\vec{e},\vec{n}) := \#\{i|n_i = 1,\ e_i = (0,1) \text{ or } (1,1)\}.$$

Theorem 3 *Assume that the binary matrix M_p satisfies the condition (1). If the distribution $\mathcal{P}_{\vec{n}}$ takes positive probabilities only in the set $\{\vec{e}|t(\vec{e},\vec{n}) = t\}$, then we obtain*

$$\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min|M_p} \le 2^{K^1(\vec{n})\overline{h}(\frac{t}{K^1(\vec{n})})+K^2(\vec{n})-m}.$$

Further, if the stochastic behavior of the random variable $t = t(\vec{e},\vec{n})$ on the distribution $\mathcal{P}_{\vec{n}}$ is described by the distribution $p(\frac{t}{K^1(\vec{n})})$, then the inequality

$$\begin{aligned}&\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min|M_p}\\ \le& \sum_{t=0}^{K^1(\vec{n})} p(\frac{t}{K^1(\vec{n})}) \min\left\{2^{K^1(\vec{n})\overline{h}(\frac{k}{K^1(\vec{n})})+K^2(\vec{n})-m}, 1\right\}\end{aligned}$$

holds. That is, the upper bound can be characterized by $\vec{K}(\vec{n})$ and t.

For a proof, see Appendix D.

C. Security of known channel: dark count case

Next, we take into account the effect of dark count in the detector. In this case, in order to characterize the presence or the absence of dark count, we add c or d to the label n_i of the input system. That is, the label n_i is chosen among $(0,c),(0,d),(1,c),(1,d),(2,+,c),(2,+,d),(2,\times,c),(2,\times,d),\ldots$ etc, where $(*,d)$ expresses dark count and $(*,c)$ does the normal count. Then, we can classify detected pulses to the following six parts:

$o = 0$, $J^0(\vec{n})$: The number of detected pulses except for dark count whose initial (Alice's) state is the vacuum state.

$o = 1$, $J^1(\vec{n})$: The number of detected pulses except for dark count whose initial (Alice's) state is the single-photon state.

$o = 2$, $J^2(\vec{n})$: The number of detected pulses except for dark count whose initial (Alice's) state is the multi-photon state.

$o = 3$, $J^3(\vec{n})$: The number of pulses detected by dark count whose initial (Alice's) state is the vacuum state.

$o = 4$, $J^4(\vec{n})$: The number of pulses detected by dark count whose initial (Alice's) state is the single-photon state.

$o = 5$, $J^5(\vec{n})$: The number of detected pulses except for dark count whose initial (Alice's) state is the multi-photon state.

Now, we consider the following protocol: First, Alice sends the random number with the $+$ basis via $\sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}_{e_i|n_i})(\rho)$, where for the dark counts $n_i = (*,d)$, e_i takes only d and the map $\mathcal{E}_{d|n_i}$ is given by $\mathcal{E}_{d|n_i}(\rho) = \frac{1}{2}(|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|)$. Second, they apply the forward or reverse error correction by the code $\mathrm{Im}\, M_e$, and finally perform privacy amplification by M_p, where M_p is assumed to satisfy (1). In this case, we obtain the same argument as Theorem 2. Thus, in order to discuss the security, we need to characterize the average error probability in the $\times$ basis.

Now, we consider the forward error correction case. In the event $o = 0, 3$, Eve cannot obtain any information of Alice's raw key. Also, in the event $o = 2, 4, 5$, Eve can obtain all information of Alice's raw key. Thus, our situation is the same as the case of $K^1 = J^1$ and $K^2 = J^2 + J^4 + J^5$ of Theorem 3. Similar to the above subsection, we define the average error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min,\to|M_p}$ of the code $(M_e(\mathrm{Ker}\, M_p))^\perp/(\mathrm{Im}\, M_e)^\perp$ concerning the $\times$ basis with the channel $\sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}'_{e_i|n_i})(\rho)$, where we define the map $\mathcal{E}'_{e_i|n_i}$ for dark count $n_i = (*, d)$ as follows: $\mathcal{E}'_{d|(*,d)}$ is the same as $\mathcal{E}_{d|*}$ and

$$\mathcal{E}'_{d|(0,d)}(\rho) := \frac{1}{2}(\mathsf{X}^0\rho(\mathsf{X}^0)^\dagger + \mathsf{X}^1\rho(\mathsf{X}^1)^\dagger),$$
$$\mathcal{E}'_{d|(*,d)}(\rho) := \frac{1}{2}(\mathsf{Z}^0\rho(\mathsf{Z}^0)^\dagger + \mathsf{Z}^1\rho(\mathsf{Z}^1)^\dagger)$$

for $* \neq 0$. The distribution $p(\frac{t}{J^1(\vec{n})})$ is defined as the distribution describing the random variable $t = t(\vec{e}, \vec{n})$ under the distribution $\mathcal{P}_{\vec{n}}$, and define $t(\vec{e}, \vec{n})$ by

$$t(\vec{e}, \vec{n}) := \#\{i | n_i = (1, c),\ e_i = (0, 1) \text{ or } (1, 1)\}.$$

Then, we obtain

$$\begin{aligned} &\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min,\to|M_p} \\ \le &\sum_{t=0}^{J^1} p(\frac{t}{J^1}) \min\left\{2^{J^1 \overline{h}(\frac{t}{J^1}) + J^2 + J^4 + J^5 - m}, 1\right\}. \end{aligned} \tag{14}$$

Next, we consider the reverse error correction case. We assume that the bits detected by dark count cannot be controlled by Eve. That is, in the event $o = 3, 4, 5$, Eve cannot obtain any information of Bob's raw key. Also, in the event $o = 0, 2$, Eve can obtain all information of Bob's raw key. (In the case of $o = 0$, Eve can obtain Bob's information by the following. Eve generates an entangled pair, and sends Bob a part of it. After announcing the basis, Eve measures the remaining part based on the correct basis.) Hence, our situation is the same as the case of $K^1 = J^1$ and $K^2 = J^0 + J^2$ of Theorem 3. Similar to the above subsection, we define the average error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftarrow|M_p}$ of the code $(M_e(\mathrm{Ker}\, M_p))^\perp/(\mathrm{Im}\, M_e)^\perp$ concerning the $\times$ basis with the channel $\sum_{\vec{e}} \mathcal{P}_{\vec{n}}(\vec{e})(\otimes_{i=1}^N \mathcal{E}'_{e_i|n_i})$, where we define the map $\mathcal{E}'_{e_i|n_i}$ as follows:

$$\mathcal{E}'_{d|(\cdot,d)}(\rho) := \frac{1}{2}(\mathsf{X}^0\rho(\mathsf{X}^0)^\dagger + \mathsf{X}^1\rho(\mathsf{X}^1)^\dagger),$$
$$\mathcal{E}'_{e_i|(1,c)}(\rho) := \mathsf{W}^{e_i}\rho(\mathsf{W}^{e_i})^\dagger,$$
$$\mathcal{E}'_{e_i|(*,c)}(\rho) := \frac{1}{2}(\mathsf{Z}^0\rho(\mathsf{Z}^0)^\dagger + \mathsf{Z}^1\rho(\mathsf{Z}^1)^\dagger)$$

for $* \neq 1$.

Note that the definition of $\mathcal{E}'_{e_i|n_i}$ for dark count $n_i = (*, d)$ is different from the forward case. Here $x = \leftarrow$ expresses the reverse case. Then, we obtain

$$\begin{aligned} &\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftarrow|M_p} \\ \le &\sum_{t=0}^{J^1} p(\frac{t}{J^1}) \min\left\{2^{J^1 \overline{h}(\frac{t}{J^1}) + J^0 + J^2 - m}, 1\right\}. \end{aligned} \tag{15}$$

D. Security of unknown channel: dark count case

Now, we back to the original setting. Since the numbers $J^0, \ldots, J^5$ and the ratio $r^1 := \frac{t}{J^1}$ are unknown, the size of sacrifice bits is chosen as the function $m(\mathcal{D}_i, \mathcal{D}_e)$ of the random variable $\mathcal{D}_e$. For simplicity, we abbreviate $m(\mathcal{D}_i, \mathcal{D}_e)$ and $\eta(\frac{H_{i_0+k}}{E_{i_0+k} - N})$ to m and η.

Now, we give general security formulas for the given function m of $\mathcal{D}_e$. The random variable $\vec{n}$ is known by Eve, but cannot be decided by Eve. Hence, Eve's information is measured by the conditional expectations $I^{\mathcal{P}}_{E|M_p,\mathcal{D}_e,\mathrm{POS}}$ of $I^{\mathcal{P}_{\vec{n}}}_{E|M_p}$ concerning the random variable $\vec{n}$ when the random variables $M_p, \mathcal{D}_e$, and POS are fixed, where POS is the random variable describing the position of the check bits and each kinds of pulses. We define $P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}$ as the conditional expectations of $P^{\mathcal{P}_{\vec{n}}}_{ph,min,x|M_p,\vec{J}}$ concerning $\vec{n}$ when the random variables $\mathcal{D}_e$, and POS are fixed. Then, we obtain

$$\begin{aligned} &\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} I^{\mathcal{P}}_{E|M_p,\mathcal{D}_e,\mathrm{POS}} \\ \le &\mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}} \overline{h}(P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}}) \\ &\quad + \mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}}(N\eta - m) P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}} \\ \le &P^{\mathcal{P}}_{ph,av,x}(\overline{N} + 1 - \log P^{\mathcal{P}}_{ph,av,x}), \end{aligned} \tag{16}$$

where $\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}}$ ($\mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}}$) denotes the expectation concerning the random variables $M_p, \mathcal{D}_e$, and POS, ($\mathcal{D}_e$, and POS). The inequality (16) is proved in Appendix E. Hence, Eve's information per one bit can be evaluated as

$$\begin{aligned} \mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}} \frac{I^{\mathcal{P}}_{E|\mathcal{D}_e,\mathrm{POS}}}{N\eta - m} \le &\mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}} \frac{\overline{h}(P^{\mathcal{P}}_{ph|\mathcal{D}_e,\mathrm{POS}})}{N\eta - m} + P^{\mathcal{P}}_{ph,av,x} \\ \le &\frac{\overline{h}(P^{\mathcal{P}}_{ph,av,x})}{\underline{N}} + P^{\mathcal{P}}_{ph,av,x}. \end{aligned}$$

Similarly, Eve's state can be given as the conditional average Eve's state $\rho^E_{[Z]|M_p,\mathcal{D}_e,\mathrm{POS}}$ with the final key $[Z]$ when the random variables $M_p, \mathcal{D}_e$, and POS are fixed.

Then,

$$\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \min_{[Z]\neq[Z']} F(\rho^E_{[Z]},\rho^E_{[Z']|M_p,\mathcal{D}_e,\mathrm{POS}})$$
$$\geq 1-2P^{\mathcal{P}}_{ph,av,x}$$
$$\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \max_{[Z]\neq[Z']} \|\rho^E_{[Z]|M_p,\mathcal{D}_e,\mathrm{POS}}-\rho^E_{[Z']|M_p,\mathcal{D}_e,\mathrm{POS}}\|_1$$
$$\leq 4P^{\mathcal{P}}_{ph,av,x}$$
$$\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \min_{[Z]} F(\rho^E_{[Z]|M_p,\mathcal{D}_e,\mathrm{POS}},\overline{\rho}^E_{M_p,\mathcal{D}_e,\mathrm{POS}})$$
$$\geq 1-P^{\mathcal{P}}_{ph,av,x}$$
$$\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \max_{[Z]} \|\rho^E_{[Z]|M_p,\mathcal{D}_e,\mathrm{POS}}-\overline{\rho}^E_{M_p,\mathcal{D}_e,\mathrm{POS}}\|_1$$
$$\leq 2P^{\mathcal{P}}_{ph,av,x},$$

where $\overline{\rho}^E_{M_p,\mathcal{D}_e,\mathrm{POS}}$ is the average state of $\rho^E_{[Z]|M_p,\mathcal{D}_e,\mathrm{POS}}$ concerning $[Z]$, and $P^{\mathcal{P}}_{ph,av,x} := \mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}} P^{\mathcal{P}}_{ph|\mathcal{D}_e,\mathrm{POS}}$. We can evaluate the probability $P^{\mathcal{P}}_{succ,x|M_p,\mathcal{D}_e,\mathrm{POS}}$ that Eve successfully detects the final key $[Z]$:

$$\begin{aligned}
&\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} P^{\mathcal{P}}_{succ,x|M_p} \\
\leq &\mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \Bigg(\sqrt{P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}}\sqrt{1-2^{-(\eta N-m)}} \\
&\quad + \sqrt{1-P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}}\sqrt{2^{-(\eta N-m)}}\Bigg)^2 \qquad (17)\\
\leq &\left(\sqrt{P^{\mathcal{P}}_{ph,av,x}}\sqrt{1-2^{-\underline{N}}}+\sqrt{1-P^{\mathcal{P}}_{ph,av,x}}\sqrt{2^{-\underline{N}}}\right)^2, \qquad (18)
\end{aligned}$$

where (18) follows from the concavity of left hand side of (10).

In order to guarantee the security, it is sufficient to show that the probability $P^{\mathcal{P}}_{ph,av,x}$ is quite small for any $\mathcal{P}$. Since the quantity $P^{\mathcal{P}}_{ph,av,x}$ has the linear form concerning $\mathcal{P}$, it is enough to treat $P^{\mathcal{P}}_{ph,av,x}$ when $\mathcal{P}$ is an extremal point. That is, the relation

$$\max_{\mathcal{P}:\text{ any conditional distribution}} P^{\mathcal{P}}_{ph,av,x} = \max_{\mathcal{P}\in\mathcal{EP}} P^{\mathcal{P}}_{ph,av,x}$$

holds, where $\mathcal{EP}$ is the set of extremal points concerning the set of conditional distributions. From (14) and (15), these values are evaluated as follows.

$$\begin{aligned}
&\max_{\mathcal{P}\in\mathcal{EP}} P^{\mathcal{P}}_{ph,av,\rightarrow} \\
\leq &\max_{\mathcal{P}\in\mathcal{EP}} \mathrm{E}^{\mathcal{P}}_{\vec{J},t,\mathcal{D}_e,\mathrm{POS}} \min\left\{2^{J^1\overline{h}(\frac{t}{J^1})+J^2+J^4+J^5-m},1\right\} \qquad (19)\\
&\max_{\mathcal{P}\in\mathcal{EP}} P^{\mathcal{P}}_{ph,av,\leftarrow} \\
\leq &\max_{\mathcal{P}\in\mathcal{EP}} \mathrm{E}^{\mathcal{P}}_{\vec{J},t,\mathcal{D}_e,\mathrm{POS}} \min\left\{2^{J^1\overline{h}(\frac{t}{J^1})+J^0+J^2-m},1\right\}. \qquad (20)
\end{aligned}$$

where t is the number of errors of the $\times$ basis in the event of $o=1$. Here, we have to treat the conditional expectation concerning $\vec{J}$ and t even if the other random variable $\mathcal{D}_e$ is fixed. Hence, our purpose is choosing the size m of sacrifice bits based on the information $\mathcal{D}_i$ and $\mathcal{D}_e$.

Since Alice chooses the positions of each kinds of pulses and check bits randomly, as is discussed in Hayashi et al. [15], the stochastic behavior of $\mathcal{D}_e$ is given by hypergeometric distribution in the case of any extremal point $\mathcal{P}$. In order to guarantee the security with the finite-length code, we have to calculate (19) and (20) for the specific function $m(\mathcal{D}_i,\mathcal{D}_e)$. Since this task needs a large amount of calculation due to a large number of random variables, we treat it in another paper [15].

In the beginning of this section, we assume that the states $\{|n,0\rangle\langle n,0|, |0,n\rangle\langle 0,n|, |n,0,\times\rangle\langle n,0,\times|, |0,n,\times\rangle\langle 0,n,\times|\}_{n\geq 2}$ can be distinguished by Eve. Now, instead of the above states, we focus on the other set of states $\{\rho_i^{\uparrow,+} := \sum_n s^i_n|n,0\rangle\langle n,0|, \rho_i^{\downarrow,+} := \sum_n s^i_n|0,n\rangle\langle 0,n|, \rho_i^{\uparrow,\times} := \sum_n s^i_n|n,0,\times\rangle\langle n,0,\times|, \rho_i^{\downarrow,\times} := \sum_n s^i_n|0,n,\times\rangle\langle 0,n,\times|\}_{n\geq 2}$ which can describe all sent pulses by the convex combination of theirselves with the states $|0,0\rangle\langle 0,0|$, $|1,0\rangle\langle 1,0|$, $|0,1\rangle\langle 0,1|$, $|1,0,\times\rangle\langle 1,0,\times|$, $|0,1,\times\rangle\langle 0,1,\times|$. Then, we can assume so stronger ability of Eve that Eve can distinguish all states of $\{\rho_i^{\uparrow,+},\rho_i^{\downarrow,+},\rho_i^{\uparrow,\times},\rho_i^{\downarrow,\times}\}_{n\geq 2}$. In this case, we obtain the same argument as this section with replacing the former set by the latter set. The construction of s^i_n in the case of the phase-randomized coherent light is given in Hayashi [16].

E. Security with two-way error correction

Here, we should remark that the effects of dark counts and the vacuum states are helpful only when the error correction is one-way. If we apply a careless two-way error correction, these effects are not so helpful. That is, Eve has a possibility to access the information in the events $o=0,2,3,4,5$. The main point of the two-way error correction is the following: Consider the case where a reverse error correction is applied after a forward error correction. In this case, the second error correction depends on (a prat of) Bob's syndrome. That is, he has to announce (a part of) his syndrome. Now, consider an extremal case, i.e., the case where Bob announces all of his syndrome. This case is equivalent with the case where Bob announces his syndrome after Alice transmits her information via a Pauli channel with the + basis.

In the single-photon case, as is discussed in Appendix B, Eve's information contains all information concerning the flip action X on the + basis, which includes Bob's syndrome. Hence, this information it is useless for Eve in the single-photon case. However, it allows Eve to access the information in the events $o=0,2,4,5$ in the imperfect photon case. Eve knows the parts $o=2,4,5$ concerning Alice's bits Z after the forward error correction by the code $C\subset\mathbf{F}_2^N$. She also knows the parts $o=0,2$ con-

cerning Bob's bits Z' after the forward error correction by the code C using Bob's syndrome. The channels in other parts $o = 1, 3$ can be regarded as the single-photon case with the channels:

$$\mathcal{E}'_{e_i|(1,c)}(\rho) = \mathsf{W}^{e_i}\rho(\mathsf{W}^{e_i})^\dagger$$
$$\mathcal{E}'_{d|(0,d)}(\rho) = \frac{1}{2}(\mathsf{Z}^0\rho(\mathsf{Z}^0)^\dagger + \mathsf{Z}^1\rho(\mathsf{Z}^1)^\dagger).$$

Suppose that Bob can perfectly correct the error, i.e., his bits Z' is equal to hers Z. Eve knows the parts $o = 0, 2, 4, 5$ concerning Z. Now, we focus on the subcode $C' \subset \mathbf{F}_2^{J^1+J^3}$ defined by

$$C' = \{x \in \mathbf{F}_2^{J^1+J^3} | (x, \vec{0}_{N-(J^1+J^3)}) \in C\},$$

where $\vec{0}_{N-(J^1+J^3)}$ is the 0 vector in the composite system of the parts $o = 0, 2, 4, 5$. Then, Eve's state is equal to that in the case where Alice sends her information with the code C' via the $J^1 + J^3$-qubits channel $\sum_{e_i} P_{\vec{n}}(\vec{e}) \otimes_{i:n_i=(1,c),(0,d)} \mathcal{E}'_{e_i|n_i}$. Therefore, our situation is the same as the case of $K^1 = J^1$ and $K^2 = J^0 + J^2 + J^4 + J^5$ of Theorem 3.

Now, we proceed to the general case of two-way error correction, in which the final classical error correction code C^u is chosen with the probability $p(u)$, i.e., Alice decides the i-th code C_i depending on the $i-1$ syndromes of Bob, inductively. We define the average error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftrightarrow|M_p,u}$ and the distribution $p_u(\frac{t}{J^1})$ concerning the random variable $\frac{t}{J^1}$ when the classical error correction code C^u is chosen, where t is defined similar to subsection III B. The relation

$$\sum_u p(u)p_u(\frac{t}{J^1}) = p(\frac{t}{J^1}) \tag{21}$$

holds. Applying Theorem 3 in the case of $K^1 = J^1$ and $K^2 = J^0 + J^2 + J^4 + J^5$, we obtain

$$\begin{aligned}&\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftrightarrow|M_p,u}\\ \le &\sum_{t=0}^{J^1} p_u(\frac{t}{J^1}) \min\left\{2^{J^1\overline{h}(\frac{t}{J^1})+J^0+J^2+J^4+J^5-m}, 1\right\}.\end{aligned} \tag{22}$$

Thus, from (21) and (22), the average error probability $P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftrightarrow|M_p}$ satisfies that

$$\begin{aligned}&\mathrm{E}_{M_p} P^{\mathcal{P}_{\vec{n}}}_{ph,min,\leftrightarrow|M_p}\\ \le &\sum_{t=0}^{J^1} p(\frac{t}{J^1}) \min\left\{2^{J^1\overline{h}(\frac{t}{J^1})+J^0+J^2+J^4+J^5-m}, 1\right\}.\end{aligned} \tag{23}$$

Thus, we can derive the same argument as subsection III D. Here, the choice of the sacrifice bit size m depends only on the data $\mathcal{D}_i$ and $\mathcal{D}_e$. If we choose the sacrifice bit size m using information u, there is a possibility to improve the above evaluation.

IV. ASYMPTOTIC KEY GENERATION RATE

A. Asymptotic key generation rate with dark count effect

From the discussion of the precious section, if we choose the number of sacrifice bits m as a larger number than $J^1\overline{h}(r^1) + J^2 + J^4 + J^5$ in the forward case, our final key is asymptotically secure. Hence, we call $J^1\overline{h}(r^1) + J^2 + J^4 + J^5 = N - J^1(1 - \overline{h}(r^1)) - (J^0 + J^3)$ the initial Eve's information in the forward case. Also, $J^1\overline{h}(r^1) + J^0 + J^2 = N - J^1(1 - \overline{h}(r^1)) - (J^3 + J^4 + J^5)$ is called the initial Eve's information in the reverse case. Thus, the asymptotic key generation (AKG) rates for the detected pulses of the forward and reverse cases are equal to

$$\frac{J^1(1 - \overline{h}(r^1)) + J^0 + J^3}{N} - (1 - \eta(s_{\nu,+})) \tag{24}$$
$$\frac{J^1(1 - \overline{h}(r^1)) + J^3 + J^4 + J^5}{N} - (1 - \eta(s_{\nu,+})), \tag{25}$$

respectively, when $\eta(s_{\nu,+})$ is the coding rate of the classical error correction code, where $N := \sum_{i=0}^5 J_i$ and $s_{\nu,+}$ is the average error probability of the detected pulses.

In the asymptotic case, $\frac{J^3+J^4+J^5}{N}$ and $\frac{J^0+J^3}{N}$ converge to p_D and $\nu(0)p_0$ in probability, respectively, where p_0 is the counting rate of the vacuum pulse and p_D is the rate of the dark counts among sent pulses. Thus, when our pulse is generated by the distribution ν, the initial Eve's informations in the forward and reverse cases are equal to

$$N(1 - \frac{\nu(1)q^1(1 - \overline{h}(r^1))}{p_{\nu,+}} - \frac{\nu(0)p_0}{p_{\nu,+}}) \tag{26}$$
$$N(1 - \frac{\nu(1)q^1(1 - \overline{h}(r^1))}{p_{\nu,+}} - \frac{p_D}{p_{\nu,+}}), \tag{27}$$

respectively, where $p_{\nu,+}$ is the counting rate of the pulse with the $+$ basis generated by the distribution ν, q^1 is the counting rate of the single-photon states except for dark counts, and r^1 is the error rate of the $\times$ basis among the single-photon states detected except for dark counts. Hence, two important rates q^1 and r^1 are needed to be estimated.

By taking into account the counting rate $p_{\nu,+}$, the AKG rates for the sent pulses of the forward and reverse cases are equal to

$$I_\rightarrow := \frac{\nu(1)q^1(1 - \overline{h}(r^1)) + \nu(0)p_0 - p_{\nu,+}(1 - \eta(s_{\nu,+}))}{2} \tag{28}$$
$$I_\leftarrow := \frac{\nu(1)q^1(1 - \overline{h}(r^1)) + p_D - p_{\nu,+}(1 - \eta(s_{\nu,+}))}{2}, \tag{29}$$

respectively, where $s_{\nu,+}$ is the error rate of pulses generated with the distribution ν in the $+$ basis. These rates are equal to those conjectured by BBL[14]. Hence, the

difference between $\frac{\nu(0)p_0}{2}$ and $\frac{p_D}{2}$ gives those of the forward and reverse cases.

By applying GLLP[12]-ILM[7] formulas, the AKG rate is equal to

$$I_{GLLP-ILM} := \frac{1}{2}\left(\nu(1)\overline{q^1}(1-\overline{h}(\overline{r^1})) - p_{\nu,+}(1-\eta(s_{\nu,+}))\right),$$

where $\overline{q^1}$ is the rate of all detected single-photon states (containing states detected by dark counts), and $\overline{r^1}$ is the error rate among all detected single-photon states in the $\times$ basis [4]. These are calculated as

$$\begin{aligned}\overline{q^1} &= q^1 + p_D \\ \overline{r^1} &= \frac{r^1 q^1 + \frac{1}{2}p_D}{q^1 + p_D}.\end{aligned}$$

If we do not take into account the effect of dark counts, the AKG rates of the forward and reverse cases are calculated to

$$\begin{aligned}\overline{I}_{\rightarrow} :=& \frac{\nu(1)\overline{q^1}(1-\overline{h}(\overline{r^1})) + \nu(0)p_0 - p_{\nu,+}(1-\eta(s_{\nu,+}))}{2} \\ \overline{I}_{\leftarrow} :=& \frac{\nu(1)\overline{q^1}(1-\overline{h}(\overline{r^1})) - p_{\nu,+}(1-\eta(s_{\nu,+}))}{2},\end{aligned}$$

respectively. The AKG rate $\overline{I}_{\rightarrow}$ was conjectured by Lo [13], and proved by Koashi [19] independently.

The discussion in subsection III E implies that the AKG rate

$$I_{\leftrightarrow} := \frac{\nu(1)q^1(1-\overline{h}(r^1)) + \nu(0)p_D - p_{\nu,+}(1-\eta(s_{\nu,+}))}{2}$$

can be attained by two-way error correction[21]. Assuming that the coding rate of two-way error correction is equal to that of one-way error correction, we compare these AKG rates. Since $\nu(1)q^1(1-\overline{h}(r^1)) + \nu(0)p_D \geq \nu(1)q^1(1-\overline{h}(r^1)) = \nu(1)(q^1(1-\overline{h}(r^1)) + p_D(1-\overline{h}(\frac{1}{2}))) \geq \nu(1)(q^1+p_D)(1-\overline{h}(\frac{r^1q^1+\frac{1}{2}p_D}{q^1+p_D})) = \nu(1)\overline{q^1}(1-\overline{h}(\overline{r^1}))$, we have

$$\begin{aligned}& I_{\rightarrow} \geq \overline{I}_{\rightarrow} \geq I_{GLLP-ILM} \\ & I_{\leftarrow} \geq I_{\leftrightarrow} \geq \overline{I}_{\leftarrow} \geq I_{GLLP-ILM} \\ & I_{\rightarrow} \geq I_{\leftrightarrow}.\end{aligned}$$

B. Mixture of the vacuum state and the single-photon state

First, we assume that $p_S = 0$. Now, we consider the distribution ν taking probabilities only in the vacuum state and the single-photon state. Then, q^1 and the error rate $r^1 = r^1_\times$ of the $\times$ basis can be solved from the counting rate p_0 of the vacuum states, the counting rate $p_{\nu,\times}$ of the pulses generated by ν in the $\times$ basis, and the error rate $s_{\nu,\times}$ of the same pulses as follows. Since q^1 and $r^1_\times$ satisfy the equations:

$$\begin{aligned}p_{\nu,\times} &= \nu(0)p_0 + \nu(1)(p_D + q^1) \\ s_{\nu,\times}p_{\nu,\times} &= \frac{1}{2}\nu(0)p_0 + \nu(1)(\frac{1}{2}p_D + r^1_\times q^1),\end{aligned}$$

we obtain

$$\begin{aligned}q^1 &= \frac{p_{\nu,\times} - \nu(0)p_0}{\nu(1)} - p_D \\ r^1_\times &= \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}\nu(0)p_0 - \frac{1}{2}\nu(1)p_D}{p_{\nu,\times} - \nu(0)p_0 - \nu(1)p_D}.\end{aligned}$$

Note that the counting rate $p_{\nu,+}$ of the pulses generated by ν in the $+$ basis coincides with the counting rate $p_{\nu,\times}$ of the pulses generated by ν in the $\times$ basis. In the case of $p_S \neq 0$, $r^1_\times$ can be calculated as

$$r^1_\times = \frac{\frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}\nu(0)p_0 - \frac{1}{2}\nu(1)p_D}{p_{\nu,\times} - \nu(0)p_0 - \nu(1)p_D} - p_S}{1-2p_S}.$$

This is because when $r^{1\prime}$ is the error probability among the single-photon states detected except for dark counts, the relation $r^{1\prime} = p_S(1-r^1) + (1-p_S)r^1$ holds.

C. Approximate single-photon state

Now, in the case of $p_S = 0$, we discuss the distribution ν taking probabilities not only in the vacuum state and the single-photon state but also in multi-photon states. Approximate single-photon state has this form. When we can generate pulses only with the distribution ν, we have to treat the rates $q^2_\times$ and q^2_+ of counuts except for dark counts of the multi-photon states in the $\times$ and $+$ bases and the error rates $r^2_\times$ and r^2_+ of the multi-photon states detected in the $\times$ and $+$ bases except for dark counts as unknown parameters as well as the rate q^1 of counts except for dark counts of the single-photon states and the error rates $r^1_\times$ and r^1_+ of the $\times$ basis and the $+$ basis of the single-photon states detected except for dark counts. Thus, the following equations hold:

$$\begin{aligned}& p_{\nu,\times} \\ =& \nu(0)p_0 + \nu(1)(p_D + q^1) + \nu(2)(p_D + q^2_\times) \qquad (30) \\ & p_{\nu,+} \\ =& \nu(0)p_0 + \nu(1)(p_D + q^1) + \nu(2)(p_D + q^2_+) \qquad (31) \\ & s_{\nu,\times}p_{\nu,\times} \\ =& \frac{1}{2}\nu(0)p_0 + \nu(1)(\frac{1}{2}p_D + r^1_\times q^1) + \nu(2)(\frac{1}{2}p_D + r^2_\times q^2_\times) \qquad (32) \\ & s_{\nu,+}p_{\nu,+} \\ =& \frac{1}{2}\nu(0)p_0 + \nu(1)(\frac{1}{2}p_D + r^1_+ q^1) + \nu(2)(\frac{1}{2}p_D + r^2_+ q^2_+), \qquad (33)\end{aligned}$$

where $q^1, q^2_\times$, and q^2_+ belong to the interval $[0, 1-p_D]$, and $r^1_\times, r^1_+, r^2_\times$, and r^2_+ do to the interval $[0,1]$. The AKG rate is characterized by the minimum value of $q^1(1-\overline{h}(r^1_\times))$ with these conditions. Since it is difficult to calculate this minimum, we treat the symmetric case, i.e., the case:

$$p_{\nu,\times} = p_{\nu,+}, \quad s_{\nu,\times} = s_{\nu,+}. \tag{34}$$

Then, the minimum $q^1_{\min}$ of q^1 and the maximum $r^1_{\max}$ of $r^1_\times$ are given as

$$\begin{aligned} q^1_{\min} &= \frac{p_{\nu,\times} - p_0\nu(0) - \nu(2)}{\nu(1)} - p_D \\ r^1_{\max} &= \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}. \end{aligned}$$

The minimum $q^1_{\min}$ and and the maximum $r^1_{\max}$ are realized simultaneously when $q^2_\times = 1-p_D, r^2_\times = 0$. The minimum of $q^1(1-\overline{h}(r^1_\times))$ is equal to $q^1_{\min}(1-\overline{h}(r^1_{\max}))$.

Next, we consider how much AKG rate can be improved when we send pulses generated by different distributions. For this purpose, we focus on the maximum $q^1_{\max}$ of q^1 and the minimum $r^1_{\max}$ of $r^1_\times$, which are calculated as

$$\begin{aligned} q^1_{\max} &= \frac{p_{\nu,\times} - p_0\nu(0) - \nu(2)p_D}{\nu(1)} - p_D \\ r^1_{\min} \le \tilde{r}^1_{\min} &:= \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2) - (1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)p_D}. \end{aligned}$$

The difference between the maximum and the minimum are given as

$$\begin{aligned} & q^1_{\max} - q^1_{\min} = \frac{\nu(2)(1-p_D)}{\nu(1)} \\ & r^1_{\max} - r^1_{\min} \le r^1_{\max} - \tilde{r}^1_{\min} \\ \le & r^1_{\max} - \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2) - (1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\left(1 - \frac{\nu(2)(1-p_D)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\right) \\ = & \frac{(1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\left(1 + \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2) - (1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\right), \end{aligned}$$

where the inequality $\frac{a}{b+x} = \frac{a}{b}\frac{1}{1+\frac{x}{b}} \ge \frac{a}{b}(1-\frac{x}{b})$ is applied in the case of $a = s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2) - (1-p_D)\nu(2)$, $b = p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)$, and $x = \nu(2)(1-p_D)$. Hence, when these differences are small relatively with $q_{\min}$ and $r_{\max}$, the AKG rate cannot be improved so much even though we send pulses generated by different distributions. For example, $(1-p_D)\nu(2)$ is small enough when the generated pulse is close enough to the single-photon.

When the symmetric assumption (34) does not hold, the conditions (31) and (33) are added with the conditions (30) and (32). The maximums $q^1_{\max}$ and $r^1_{\max}$ become small, and the minimums $q^1_{\min}$ and $r^1_{\min}$ become large. Hence, the following relations also hold even in the non-symmetric case:

$$\begin{aligned} q^1_{\max} - q^1_{\min} \le & \frac{\nu(2)(1-p_D)}{\nu(1)} \\ r^1_{\max} - r^1_{\min} \le & \frac{(1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\left(1 + \frac{s_{\nu,\times}p_{\nu,\times} - \frac{1}{2}p_0\nu(0) - \frac{1}{2}p_D\nu(1) - \frac{1}{2}p_D\nu(2) - (1-p_D)\nu(2)}{p_{\nu,\times} - p_0\nu(0) - p_D\nu(1) - \nu(2)}\right). \end{aligned}$$

V. DETAIL ANALYSIS ON EVE'S ATTACK

A. Reduction to three-dimensional outcome channel

We prove that any Eve's attack can be reduced by the attack discussed in Section III. Of course, in the following discussion contains the case when the frame of Alice

does not coincide with that of Bob. Since Alice sends N pulses and Bob receives N pulses, Eve's operation can be described by a CP-TP map $\mathcal{E}_N$ from the system $\mathcal{H}^{\otimes N}$ to the system $\mathcal{H}^{\otimes N}$. This description contains the loss of the communication channel. Even if the detector has the loss, if the loss does not depend on the measurement basis, the security is guaranteed by our discussion.

In order to reduce the output system $\mathcal{H}$ to three-dimensional system $\mathcal{H}_0 \oplus \mathcal{H}_1$, we modify our protocol as follows: In the measurement with the + basis, Bob performs the measurement $\{|n,m\rangle\}_{n,m}$. When $(0,0)$ is measured, he decides his final outcome to be $\emptyset$. When (n,m) is measured, he does his final outcome to be 0 with the probability $\frac{n}{n+m}$, and 1 with the probability $\frac{m}{n+m}$. This POVM with three outcomes is denoted by $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$ on the system $\mathcal{H}$. First, we discuss the security based on the POVM $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$, and after this discussion, we treat the security with the POVM $\{M_\emptyset, M_0, M_1\}$, which is given in Section II.

The stochastic behavior of the outcome of the POVM $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$ is described by the POVM $\{M'_\emptyset := |0\rangle\langle 0|, M'_0 := |1,0\rangle\langle 1,0|, M'_1 := |0,1\rangle\langle 0,1|\}$ on the system $\mathcal{H}_0 \oplus \mathcal{H}_1$ and the TP-CP map $\mathcal{E}$, which are defined as

$$\mathrm{Tr}\, \tilde{M}_i \rho = \mathrm{Tr}\, e'_i \mathcal{E}(\rho)$$

$$\mathcal{E}(\rho) := P_{\mathcal{H}_0} \rho P_{\mathcal{H}_0} + P_{\mathcal{H}_1} \rho P_{\mathcal{H}_1} + \sum_{n=2}^{\infty} \mathcal{E}^n (P_{\mathcal{H}_n} \rho P_{\mathcal{H}_n}),$$

where the TP-CP map $\mathcal{E}^n$ from the n-photon system $\mathcal{H}_n$ (which is equal to the n-th symmetric space) to the system $\mathcal{H}_1$ is defined by embedding the state ρ on the n-th symmetric space $\mathcal{H}_n$ into n-tensor product system $\mathcal{H}_1^{\otimes n}$ as

$$\mathcal{E}^n(\rho) := \mathrm{Tr}_{2,\ldots,n}\, \rho,$$

where $\mathrm{Tr}_{2,\ldots,n}$ means taking partial trace concerning 2-th - n-th subsystems.

Hence, by extending the system under Eve's control, the environment of the TP-CP map $\mathcal{E}$ can be regarded to be under Eve's control. The channel from Alice to Bob can be described as the TP-CP map $\mathcal{E}^{\otimes N} \circ \mathcal{E}_N$ from the system system $\mathcal{H}^{\otimes N}$ to the system $(\mathcal{H}_0 \oplus \mathcal{H}_1)^{\otimes N}$. We also assume a stronger ability of Eve, i.e., all states $|n,0\rangle$, $|0,n\rangle$ $|n,0,\times\rangle$, $|0,n,\times\rangle$ can be distinguished by Eve for $n \ge 2$. Then, each pulse can be described as a state on the system $\mathcal{H}' := \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus (\oplus_{n=2}^{\infty} \mathcal{H}_{n,+}) \oplus (\oplus_{n=2}^{\infty} \mathcal{H}_{n,\times})$, where the space $\mathcal{H}_{n,+}$ is spanned by $\{|n,0\rangle, |0,n\rangle\}$, and the space $\mathcal{H}_{n,\times}$ is spanned by $\{|n,0,\times\rangle, |0,n,\times\rangle\}$. Then, the channel $\mathcal{E}^{\otimes N} \circ \mathcal{E}_N$ from Alice to Bob can be regarded as a TP-CP map from the system ${\mathcal{H}'}^{\otimes N}$ to the system $(\mathcal{H}_0 \oplus \mathcal{H}_1)^{\otimes N}$, which is denoted by $\mathcal{E}'_N$ in the following.

Now, we focus on following two pinching maps:

$$\begin{aligned}
\mathcal{E}_a(\rho) :=& P_{\mathcal{H}_1} \rho P_{\mathcal{H}_1} + |0,0\rangle\langle 0,0|\rho|0,0\rangle\langle 0,0| \\
&+ \sum_{n=2}^{\infty} |n,0\rangle\langle n,0|\rho|n,0\rangle\langle n,0| \\
&+ |0,n\rangle\langle 0,n|\rho|0,n\rangle\langle 0,n| \\
&+ |n,0,\times\rangle\langle n,0,\times|\rho|n,0,\times\rangle\langle n,0,\times| \\
&+ |0,n,\times\rangle\langle 0,n,\times|\rho|0,n,\times\rangle\langle 0,n,\times|, \\
\mathcal{E}_b(\rho) :=& P_{\mathcal{H}_0} \rho P_{\mathcal{H}_0} + P_{\mathcal{H}_1} \rho P_{\mathcal{H}_1}.
\end{aligned}$$

By extending the system controlled by Eve, the channel from Alice to Bob can be regarded as a TP-CP map $\mathcal{E}_N := \mathcal{E}_b \circ \mathcal{E}'_N \circ \mathcal{E}_a$ from the system ${\mathcal{H}'}^{\otimes N}$ to the system $(\mathcal{H}_0 \oplus \mathcal{H}_1)^{\otimes N}$, due to the forms of the measurement by Bob and the states sent by Alice.

B. Discrete twirling

In order to define discrete twirling, we define the operators X and Z on the system system $\mathcal{H}'$ by

$$\begin{aligned}
\mathsf{X}|0\rangle &= |0\rangle \\
\mathsf{X}|\uparrow\rangle &= |\downarrow\rangle, \quad \mathsf{X}|\downarrow\rangle = |\uparrow\rangle \\
\mathsf{X}|n,0\rangle &= |0,n\rangle, \quad \mathsf{X}|0,n\rangle = |n,0\rangle \\
\mathsf{X}|n,0,\times\rangle &= |n,0,\times\rangle, \quad \mathsf{X}|0,n,\times\rangle = -|0,n,\times\rangle \\
\mathsf{Z}|0\rangle &= |0\rangle \\
\mathsf{Z}|\uparrow\rangle &= |\downarrow\rangle, \quad \mathsf{Z}|\downarrow\rangle = (-1)|\uparrow\rangle \\
\mathsf{Z}|n,0\rangle &= |n,0\rangle, \quad \mathsf{Z}|0,n\rangle = -|0,n\rangle \\
\mathsf{Z}|n,0,\times\rangle &= |0,n,\times\rangle, \quad \mathsf{Z}|0,n,\times\rangle = |n,0,\times\rangle,
\end{aligned}$$

and the operators X^x and Z^z for $x,z \in \mathbf{F}_2^N$ by

$$\mathsf{X}^x = \mathsf{X}^{x_1} \otimes \cdots \otimes \mathsf{X}^{x_N}, \quad \mathsf{Z}^z = Z^{z_1} \otimes \cdots \otimes Z^{z_N}.$$

It is known that if and only if the relation

$$(\mathsf{X}^x\mathsf{Z}^z)^\dagger \mathcal{E}_N (\mathsf{X}^x\mathsf{Z}^z \rho (\mathsf{X}^x\mathsf{Z}^z)^\dagger) \mathsf{X}^x\mathsf{Z}^z = \mathcal{E}_N(\rho) \tag{35}$$

holds for any $x,z \in \mathbf{F}_2^N$, $\mathcal{E}_N$ has the form of (2). Now, we define the discrete twirling $\mathcal{E}_N$ of the map $\mathcal{E}_N$:

$$\mathcal{E}_N(\rho) := \frac{1}{2^{2N}} \sum_{x,z \in \mathbf{F}_2^N} (\mathsf{X}^x\mathsf{Z}^z)^\dagger \mathcal{E}_N (\mathsf{X}^x\mathsf{Z}^z \rho (\mathsf{X}^x\mathsf{Z}^z)^\dagger) \mathsf{X}^x\mathsf{Z}^z. \tag{36}$$

The operation of 'discrete twirling' corresponds to the following operation: First, Alice generates two random numbers $x,z \in \mathbf{F}_2^N$, performs the operation $\mathsf{X}^x\mathsf{Z}^z$, and sends the state via the channel $\mathcal{E}_N$, and the classical information $x,z \in \mathbf{F}_2^N$ via the public channel. Next, Bob performs the inverse operation $(\mathsf{X}^x\mathsf{Z}^z)^\dagger$ to the received system. Since the TP-CP map $\mathcal{E}$ has the covariance (35), $\mathcal{E}$ has the form (2). However, when Eve's system is extended, the implemented channel $\mathcal{E}$ is not $\mathcal{E}_N$ but only $\mathcal{E}'_N$.

C. Forward case

In the forward error correction case, by taking into account the error correction operation, the raw keys can be regarded to be transmitted via the discrete twirling of the original channel. The operation concerning x in (36) is essentially realized by the following operation. After quantum communication, Alice generates the random number X and sends the classical information $Y = M_e Z + X$. Bob regards $X' - Y$ as the final raw key. On the other hand, the input and output data are not changed when the operation corresponding to z. If Alice and Bob perform the operation concerning z in (36), Eve's performance is increased only. Hence, in the forward case, the security can be evaluated by analysis on the channel (2). Note that since our error is symmetric in the form (2), the error probability can be estimated from the random variable $\mathcal{D}_e$.

When we take into account dark counts, any channel concerning pulses detected expect for dark counts, can be described by (2) in the forward case. Eve can obtain all information when the sent pulse is not in the vacuum state but is detected by dark counts while he cannot obtain any information when the sent pulse is in the vacuum state and is detected by dark count. Hence, the analysis in Section III is valid in the forward case.

D. Reverse case

We proceed to the case where all sent states are in single-photon and the reverse error correction is applied. Our protocol is given as follows. First, Alice sends the information bits X with the $+$ basis to Bob via the quantum channel Λ, and Bob obtains the bits X' by measuring the received states with the $+$ basis. Bob generates another random number Z, and sends the classical information $Y = Z + X'$ to Alice, and Alice regards $X - Y$ as the final raw keys. Now, we consider the following modified protocol: Alice generates an entangled pair and sends one part to Bob. Bob generates another quantum state with the $+$ bit basis, and performs the Bell measurement $\{M_{(0,0)}, M_{(0,1)}, M_{(1,0)}, M_{(1,1)}\}$ to the joint system of the received system and his original system, where $M_{(0,0)}, M_{(0,1)}, M_{(1,0)}, M_{(1,1)}$ is the projection corresponding to $\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle)$, $\frac{1}{\sqrt{2}}(|0\rangle|0\rangle - |1\rangle|1\rangle)$, $\frac{1}{\sqrt{2}}(|0\rangle|1\rangle + |1\rangle|0\rangle)$, $\frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$, respectively. Then, he sends his measurement value to Alice. Alice performs the inverse transformation depending on the data to her system, and measures it with the $+$ basis. Since the map from Bob's input state to the Alice's final state satisfies the covariance (35), this channel is described by the Pauli channel $\mathcal{E}(\Lambda)$ that is the discrete twirling of Λ. The latter protocol is essentially equivalent with the former protocol with the following modification: Bob sends Alice $(0,0)$ and $(0,1)$ with the probability $\frac{1}{2}$ in the case of $Y = 0$, and sends Alice $(1,0)$ and $(1,1)$ with the probability $\frac{1}{2}$ in the case of $Y = 1$. Hence, the channel from Bob to Alice in the former case can be described by the Pauli channel $\mathcal{E}(\Lambda)$. Therefore, without loss of generality, we can assume that the original map from Alice to Bob can be regarded as a Pauli channel.

Now, in order to treat the loss during the transmission, we modify the latter protocol as follows. Alice performs the two-valued measurement $\{T_0, T_1\}$ before sending one part of the entangled pair, and sends the state $\sqrt{T_1}\rho\sqrt{T_1}$ only when 1 is detected. In this case, the map from Bob's input state to the Alice's final state can be described by a Pauli channel. This protocol is essentially equivalent with the modification of the above former protocol with the following modification: Alice performs the two-valued measurement $\{T_0, T_1\}$ before sending her state, and sends the state $\sqrt{T_1}\rho\sqrt{T_1}$ only when 1 is detected. This modification is equivalent with the lossy channel case.

Next, we consider the case where the number of photons in the input state is not fixed and no dark count is detected. In this case, we assume that Eve can know Bob's measured value when n-photon state ($n \geq 2$) or the vacuum state is transmitted. It is needed only to describe the behavior of the counting rates and the error rates, which are estimated from the random variable $\mathcal{D}_i$. Thus, we can assume that the channel from Alice to Bob can be described by (2) without loss of generality.

Finally, we consider the case with dark counts. Eve cannot obtain any Bob's information for the bits detected by dark counts. Hence, Eve's information concerning this part has no relation with the channel from Alice to Bob. Any description of this part is allowed. Thus, even if the dark counts exist, the channel from Alice to Bob can be described by (2) without loss of generality.

E. Security with the original POVM

We compare the case with the measurement $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$ and that with the measurement $\{M_\emptyset, M_0, M_1\}$. The systems controlled by Eve in these two cases are identical. The error probability based on the measurement $\{M_\emptyset, M_0, M_1\}$ is larger than that based on the measurement $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$. Thus, the estimate of the error rate r^1 with the measurement $\{M_\emptyset, M_0, M_1\}$ is larger than that with the measurement $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$. Since a larger estimate of r^1 gives a larger size m of sacrifice bits, the security based on the measurement $\{\tilde{M}_\emptyset, \tilde{M}_0, \tilde{M}_1\}$ implies the security based on the original measurement $\{M_\emptyset, M_0, M_1\}$.

VI. CONCLUDING REMARKS

Applying the relation between Eve's information and phase error probability, we have derived useful upper bounds of eavesdropper's performances, i.e., eavesdropper's information and the trace norm between the Eve's states corresponding to final keys for the protocol given

in section II. Here, we have used powerful relations between the eavesdropper's performances and the phase error probability. We have also treated our channel as a TP-CP map on the two-mode bosonic system, which is the most general framework. Further, our discussion has taken into account the effect of dark counts, which forbids Eve to access Bob's bits of the pulses detected by dark count. However, our upper bounds (19) and (20) contain so many random variables that its detail numerical analysis in the case of phase-randomized coherent light is very complicated and is separately given in Hayashi et al [15]. Hence, the future problem for practical QKD is the numerical calculations of the bounds (19) and (20). On the other hand, the concrete calculation of the AKG rate is another future important topic. Also this topics is separately discussed in Hayashi [16].

We have treated the AKG rate more deeply when the generated imperfect resource is close to the single-photon. Since this resource is different from the perfect single-photon, we need the decoy method. We have compared the case where only the vacuum state is sent as a different pulse with the case where additional pulses are sent as different pulses.

This paper has treated only the binary case. However, it is easy to extend to the p-nary case, where p is a prime. In this case, we replace the two-mode bosonic system and $\mathbf{F}_2$ by the p-mode bosonic system and $\mathbf{F}_p$.

Acknowledgments

The author would like to thank Professor Hiroshi Imai of the ERATO-SORST, QCI project for support. He is grateful to Professor Hiroshi Imai, Dr. Akihisa Tomita, Dr. Tohya Hiroshima, and Mr. Jun Hasewaga for useful discussions. He thanks Dr. Francesco Buscemi for his meaningful comments. He also thanks the referee for his helpful comments.

APPENDIX A: PROOF OF THEOREM 1

Now, we prove Theorem 1. More precisely, we show the following. (1) An element $(x,y)^T \in \mathbf{F}_2^m \oplus \mathbf{F}_2^l$ belongs the image of $(\mathbf{X},I)^T$ with the probability 2^{-m} if $x \neq 0$ and $y \neq 0$. (2) it does not belong to the image of the transpose of any $l \times (l+m)$ Toeplitz matrices $(\mathbf{X},I)$ if $x \neq 0$ and $y = 0$.

Indeed, since (2) is trivial, we will show (1). For $y = (y_1,\ldots,y_m)$, we let i be the minimum index i such that $y_i \neq 0$. An element $(x,y)^T \in \mathbf{F}_2^m \oplus \mathbf{F}_2^l$ belongs to the image of $(\mathbf{X},I)^T$ if and only if $(x,y) = (\mathbf{X}^T y, y)$, which written as the following m conditions.

$$
\begin{aligned}
Y_i y_i &= x_1 - \sum_{j=i+1}^{l} Y_j y_j - x_1 \\
Y_{i+1} y_i &= x_2 - \sum_{j=i+1}^{l} Y_{j+1} y_j - x_1 \\
&\vdots \\
Y_{i+m-2} y_i &= x_{m-1} - \sum_{j=i+1}^{l} Y_{j+m-2} y_j - x_1 \\
Y_{i+m-1} y_i &= x_m - \sum_{j=i+1}^{l} Y_{j+m-1} y_j - x_1
\end{aligned}
$$

Now, the random variables $Y_{i+m},\ldots,Y_{l+m-1}$ are fixed. The m-th condition does not depend on the variables $Y_1,\ldots Y_{i+m-2}$. Hence, the m-th condition only depends on the variable Y_{i+m-1}. Therefore, the m-th condition holds with the probability $1/2$. Similarly, we can show that the $m-1$-th condition holds with the probability $1/2$ when the m-th condition holds. Thus, the l-th condition and $l-1$-th condition hold with $1/2^2$. Repeating this discussion inductively, we can conclude that the all m conditions hold with the probability 2^{-m}.

APPENDIX B: EVE'S STATES

For any Pauli channel $\mathcal{E}(\rho) = \sum_{x,z\in\mathbf{F}_2^l} P(x,z) X^x Z^z \rho (X^x Z^z)^\dagger$, the channel to Eve is given by $\mathcal{E}_E(\rho)$:

$$
\mathcal{E}_E(\rho) := \sum_{(x,z),(x',z')\in\mathbf{F}_2^{2l}} \sqrt{P(x,z)}\sqrt{P(x',z')}\,\mathrm{Tr}\, X^x Z^z \rho (X^{x'} Z^{z'})^\dagger |(x,z)\rangle\langle(x',z')|.
$$

If the input state is given by $|y\rangle$ in the $+$ basis, Eve's state can be evaluated as follows.

$$\begin{aligned}
&\mathcal{E}_E(|y\rangle\langle y|)\\
=&\sum_{x\in\mathbf{F}_2^l,z,z'\in\mathbf{F}_2^l}\sqrt{P(x,z)}\sqrt{P(x,z')}\langle y|Z^{z-z'}|y\rangle|(x,z)\rangle\langle(x,z')|\\
=&\sum_{x\in\mathbf{F}_2^l,z,z'\in\mathbf{F}_2^l}P(x)\sqrt{P(z|x)}\sqrt{P(z'|x)}(-1)^{(z-z')\cdot y}|(x,z)\rangle\langle(x,z')|\\
=&\sum_{x\in\mathbf{F}_2^l}P(x)|P,y,x\rangle\langle P,y,x|,
\end{aligned}\tag{B1}$$

where we define the vector $|P,y,x\rangle$ as

$$|P,y,x\rangle:=\sum_{z\in\mathbf{F}_2^l}(-1)^{z\cdot y}\sqrt{P(z|x)}|(x,z)\rangle.$$

That is, Eve's state is the stochastic mixture of the state $|P,y,x\rangle\langle P,y,x|$ with the probability $P(x)$. Then, Eve loses no information even if she measure the information x concerning the error of the $+$ basis.

APPENDIX C: PROOF OF THEOREM 2

The inequalities (4) and (5) are proved by Hayashi [10]. First, we prove the inequalities (6), (7), (8), and (9). As is similar to the inequalities (4) and (5), it is sufficient to show these inequalities for the corrected channel. This is because the code $\operatorname{Im} M_e/M_e(\operatorname{Ker} M_p)$ with the $+$ basis is equivalent with the code $(M_e(\operatorname{Ker} M_p))^\perp/(\operatorname{Im} M_e)^\perp$ with the $\times$ basis [10]. Thus, it is sufficient to show

$$\min_{y,y'\in\mathbf{F}_2^l}F(\mathcal{E}_E(|y\rangle\langle y|),\mathcal{E}_E(|y'\rangle\langle y'|))\ge 1-2P_{ph}\tag{C1}$$

$$\max_{y,y'\in\mathbf{F}_2^l}\|\mathcal{E}_E(|y\rangle\langle y|)-\mathcal{E}_E(|y'\rangle\langle y'|)\|_1\ge 4P_{ph}\tag{C2}$$

$$\min_{y\in\mathbf{F}_2^l}F(\mathcal{E}_E(|y\rangle\langle y|),\rho_{\mathrm{mix}})\ge 1-P_{ph}\tag{C3}$$

$$\max_{y\in\mathbf{F}_2^l}\|\mathcal{E}_E(|y\rangle\langle y|)-\rho_{\mathrm{mix}}\|_1\ge P_{ph},\tag{C4}$$

where ρ_{mix} is the maximally mixed state and the phase error probability P_{ph} is defined as

$$P_{ph}:=\sum_{x,z\in\mathbf{F}_2^l,z\neq 0}P(x,z)=\sum_{x\in\mathbf{F}_2^l}P(x,0)-1.$$

Since $\|\rho-\rho'\|_1\ge 2(1-F(\rho,\rho'))$, the inequalities (C2) and (C4) follow from (C1) and (C3). Now, we will prove (C1) and (C3). Remember the relation (B1). Since

$$\begin{aligned}
&|\langle P,y,x|P,y',x\rangle|=|\sum_{z\in\mathbf{F}_2^l}(-1)^{z\cdot(y'-y)}P(z|x)|\\
\ge&P(0|x)-|\sum_{z\in\mathbf{F}_2^l,z\neq 0}(-1)^{z\cdot(y'-y)}P(z|x)|\\
=&P(0|x)-(1-P(0|x))=2P(0|x)-1,
\end{aligned}$$

the fidelity $F(\mathcal{E}_E(|y\rangle\langle y|),\mathcal{E}_E(|y'\rangle\langle y'|))$ can be evaluated as

$$\begin{aligned}
&F(\mathcal{E}_E(|y\rangle\langle y|),\mathcal{E}_E(|y'\rangle\langle y'|))=\sum_{x\in\mathbf{F}_2^l}P(x)|\langle P,y,x|P,y',x\rangle|\\
\ge&\sum_{x\in\mathbf{F}_2^l}P(x)(2P(0|x)-1)=2\sum_{x\in\mathbf{F}_2^l}P(x,0)-1=1-2P_{ph},
\end{aligned}$$

which implies (C1). Since

$$\begin{aligned}
&\langle P,y,x|\big(\frac{1}{2^l}\sum_{y'\in\mathbf{F}_2^l}|P,y',x\rangle\langle P,y',x|\big)|P,y,x\rangle\\
=&\langle P,y,x|\big(\sum_{z\in\mathbf{F}_2^l}P(z|x)|(x,z)\rangle\langle(x,z)|\big)|P,y,x\rangle\ge P(0|x)^2,
\end{aligned}$$

we have

$$\begin{aligned}
&F\Big(\mathcal{E}_E(|y\rangle\langle y|),\sum_{y'\in\mathbf{F}_2^l}\frac{1}{2^l}\mathcal{E}_E(|y'\rangle\langle y'|)\Big)\\
=&\sum_{x\in\mathbf{F}_2^l}P(x)\sqrt{\langle P,y,x|\big(\frac{1}{2^l}\sum_{y'\in\mathbf{F}_2^l}|P,y',x\rangle\langle P,y',x|\big)|P,y,x\rangle}\\
\ge&\sum_{x\in\mathbf{F}_2^l}P(x)P(0|x)=\sum_{x\in\mathbf{F}_2^l}P(x,0)=1-P_{ph},
\end{aligned}$$

which implies (C3).

Next, we show (10). The discrimination on the set of states $\{\mathcal{E}_E(|y\rangle\langle y|)\}_{y\in\mathbf{F}_2^l}$ can be reduced to The discrimination on the set of states $\{|P,y,x\rangle\langle P,y,x|\}_{y\in\mathbf{F}_2^l}$. This set has a symmetry concerning the action of $y'\in\mathbf{F}_2^l$ as $U_{y'}:=\sum_{z\in\mathbf{F}_2^l}(-1)^{z\cdot y'}|(x,z)\rangle\langle(x,z)|$. Each one-dimensional subspace spanned by $|(x,z)\rangle$ is different representation subspace in the space spanned by $\{|(x,z)\rangle\}_{z\in\mathbf{F}_2^l}$. From Holevo[20]'s theory of covariant estimator, the minimum average error is given by the following covariant POVM $\{2^{-l}U_y|\phi\rangle\langle\phi|U_y^\dagger\}$, where $|\phi\rangle=\sum_{z\in\mathbf{F}_2^l}e^{i\theta_z}|(x,z)\rangle$. Then, the correct-decision probability is given as

$$2^{-l}|\langle P,0,x|\phi\rangle|^2=2^{-l}\Big|\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z|x)}e^{i\theta_z}\Big|^2.$$

Its maximal value is $(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z|x)}\sqrt{2^{-l}})^2$, which is attained when $e^{i\theta_z}=1$. Therefore, the optimal correct-decision probability of the set $\{\mathcal{E}_E(|y\rangle\langle y|)\}_{y\in\mathbf{F}_2^l}$ is equal to $\sum_{x\in\mathbf{F}_2^l}P(x)(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z|x)}\sqrt{2^{-l}})^2$.

Since $(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z)}\sqrt{2^{-l}})^2$ is the fidelity between the uniform distribution and the distribution P, the joint concavity of the fidelity guarantees that the concavity of $(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z)}\sqrt{2^{-l}})^2$ concerning the distribution P. Thus,

$$\begin{aligned}&\sum_{x\in\mathbf{F}_2^l}P(x)\Big(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z|x)}\sqrt{2^{-l}}\Big)^2\\ \le&\Big(\sum_{z\in\mathbf{F}_2^l}\sqrt{\sum_{x\in\mathbf{F}_2^l}P(x,z)}\sqrt{2^{-l}}\Big)^2.\end{aligned}$$

The concavity guarantees that

$$\begin{aligned}&\max_{P:P(0)=1-P_{ph}}\Big(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z)}\sqrt{2^{-l}}\Big)^2\\ =&\Big(\sum_{z\in\mathbf{F}_2^l}\sqrt{1-P_{ph}}\sqrt{2^{-l}}+(2^l-1)\sqrt{\frac{P_{ph}}{2^l-1}}\sqrt{2^{-l}}\Big)^2\\ =&\big(\sqrt{P_{ph}}\sqrt{1-2^{-l}}+\sqrt{1-P_{ph}}\sqrt{2^{-l}}\big)^2,\end{aligned}$$

which implies (10).

Applying the concavity of $(\sum_{z\in\mathbf{F}_2^l}\sqrt{P(z)}\sqrt{2^{-l}})^2$ between the distributions $(1-P_{ph},\frac{P_{ph}}{2^l-1},\dots,\frac{P_{ph}}{2^l-1})$ and $(1-P'_{ph},\frac{P'_{ph}}{2^l-1},\dots,\frac{P'_{ph}}{2^l-1})$, we obtain the concavity of $\big(\sqrt{P_{ph}}\sqrt{1-2^{-l}}+\sqrt{1-P_{ph}}\sqrt{2^{-l}}\big)^2$ concerning P_{ph}.

APPENDIX D: PROOF OF THEOREM 3

Since the condition $Z\in(M_e\operatorname{Ker}M_p)^\perp$ is equivalent with the condition $M_e^TZ\in\operatorname{Im}M_p^T$ for $Z\in\mathbf{F}_2^N$, the condition (1) is equivalent with the condition:

$$\mathrm{P}\{Z\in(M_e\operatorname{Ker}M_p)^\perp\}\le 2^{-m}\text{ for }\forall Z\in\mathbf{F}_2^N\setminus(\operatorname{Im}M_e)^\perp.$$

Hence, Theorem 3 is essentially equivalent with the following proposition, which we will prove.

Proposition 1 *Let $C_1\subset\mathbf{F}_2^N$ be an $l+m$-dimensional code. We choose an m-dimensional subcode $C_2(X)\subset C_1$ satisfying the following condition:*

$$\mathrm{P}_X\{X|x\in C_2(X)^\perp\}\le 2^{-m}\text{ for }\forall x\in\mathbf{F}_2^n\setminus C_1^\perp,\tag{D1}$$

where X is the random variable describing the stochastic behavior. Alice sends the information $C_2(X)^\perp/C_1^\perp$ via the following channel. Here, when she wants to send the information $[x]_{C_1^\perp}\in C_2(X)^\perp/C_1^\perp$, she chooses x' among $x+C_1^\perp$ with the equal probability 2^{l+m-N}. The total N bits can be divided to the following three parts:

n_0 **bits:** *no noise (0th part).*

n_1 **bits:** *at most t bits will be changed (1st part).*

n_2 **bits:** *no assumption (2nd part).*

Using the above classification, Bob recovers the original information for received information y by the following way. First, he defines the element $\Gamma(y)\in C_2(X)^\perp$ by

$$\Gamma(y):=\operatorname*{argmax}_{z\in C_2(X):\pi_0(y)=\pi_0(z)}|\pi_1(y-z)|,$$

where π_i is the projection to the above n_i bits. Next, Bob recovers the information $[\Gamma(y)]_{C_1^\perp}\in C_2(X)^\perp/C_1^\perp$.

Then, we obtain

$$\begin{aligned}&\mathrm{E}_Xp_x\{[\Gamma(y)]_{C_1^\perp}\neq[x]_{C_1^\perp}\}\\ \le&2^{n_1\overline{h}(\frac{t}{n_1})+n_2-m}\end{aligned}\tag{D2}$$

for any $[x]_{C_1^\perp}\in C_2(X)^\perp/C_1^\perp$, where p_x is the conditional distribution describing the distribution of the output y with the input x satisfying the above condition.

Proof: From the linearity,

$$\begin{aligned}p_x\{[\Gamma(y)]_{C_1^\perp}\neq[x]_{C_1^\perp}\}=&p_0\{[\Gamma(y)]_{C_1^\perp}\neq[0]_{C_1^\perp}\}\\ =&p_0\{\Gamma(y)\notin C_1^\perp\}.\end{aligned}$$

Hence, it is enough to show

$$\mathrm{E}_X\mathrm{P}_{p_0}\{\Gamma(y)\notin C_1^\perp\}\le 2^{n_1\overline{h}(\frac{t}{n_1})+n_2-m}.\tag{D3}$$

When the original massage is 0, the received signal y satisfies the conditions $\pi_0(y)=0$ and $|\pi_1(y)|\le t$, i.e., the distribution p_0 has positive probability only on the set $\mathcal{Y}:=\{y|\pi_0(y)=0,|\pi_1(y)|\le t\}$. Then,

$$\begin{aligned}&\mathrm{E}_Xp_0\{\Gamma(y)\notin C_1^\perp\}\\ =&\mathrm{E}_Xp_0\{y|\exists z\in C_2(X)^\perp\setminus C_1^\perp\text{s.t.}|\pi_1(z)-\pi_1(y)|\le|\pi_1(y)|\}\\ \le&\mathrm{E}_Xp_0\{y|\exists z\in C_2(X)^\perp\setminus C_1^\perp\text{s.t.}|\pi_1(z)-\pi_1(y)|\le t\}\\ =&\sum_{y\in\mathcal{Y}}p_0(y)\mathrm{P}_X\left\{X\,\middle|\,\begin{array}{l}\exists z\in C_2(X)^\perp\setminus C_1^\perp\\ \text{s.t.}|\pi_1(z)-\pi_1(y)|\le t\end{array}\right\}\\ \le&\sum_{y\in\mathcal{Y}}p_0(y)\sum_{z:|\pi_1(z)-\pi_1(y)|\le t}\mathrm{P}_X\{X|z\in C_2(X)^\perp\setminus C_1^\perp\}\\ \le&\sum_{y\in\mathcal{Y}}p_0(y)\sum_{z:|\pi_1(z)-\pi_1(y)|\le t}2^{-m}\qquad\text{(D4)}\\ \le&\sum_{y\in\mathcal{Y}}p_0(y)2^{n_1\overline{h}(\frac{t}{n_1})+n_2-m}=2^{n_1\overline{h}(\frac{t}{n_1})+n_2-m}.\qquad\text{(D5)}\end{aligned}$$

where the inequality (D4) follows from (D1) and the inequality (D5) does from the following inequality:

$$|\{z||\pi_1(z)-\pi_1(y)|\le t\}|\le 2^{n_1\overline{h}(\frac{t}{n_1})+n_2}.$$

Therefore, we obtain (D3). ∎

APPENDIX E: PROOF OF (16)

Since $-(1-x)\log(1-x) \le x$, we have

$$\begin{aligned}
&\overline{h}(P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}}) + \mathrm{E}^{\mathcal{P}}_{\mathcal{D}_e,\mathrm{POS}}(N\eta - m)P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}} \\
\le & - P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}} \log(P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}}) \\
& + P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}} + \overline{N} P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}} \\
= & P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}}(\overline{N} + 1 - \log P^{\mathcal{P}}_{ph,min,x|\mathcal{D}_e,\mathrm{POS}}).
\end{aligned}$$

Thus, using the concavity of the function $x \to -x\log x$, we obtain (16).

APPENDIX F: PROOF OF (18)

We will prove (18). First, the function $q \mapsto f(q) := \sqrt{1-p}\sqrt{q} + \sqrt{p}\sqrt{1-q}$ is monotone increasing for $q \le 1/2$ when $p \le 1/2$. This is because the derivative is calculated as $f'(q) = \frac{1}{2}(\sqrt{\frac{1-p}{q}} - \sqrt{\frac{p}{1-q}}) \le 0$. Thus,

$$\begin{aligned}
& \mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \Bigg(\sqrt{P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}} \sqrt{1 - 2^{-(\eta N - m)}} \\
& \quad + \sqrt{1 - P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}} \sqrt{2^{-(\eta N - m)}} \Bigg)^2 \\
\le & \mathrm{E}^{\mathcal{P}}_{M_p,\mathcal{D}_e,\mathrm{POS}} \Bigg(\sqrt{P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}} \sqrt{1 - 2^{-\underline{N}}} \\
& \quad + \sqrt{1 - P^{\mathcal{P}}_{ph,min,x|M_p,\mathcal{D}_e,\mathrm{POS}}} \sqrt{2^{-\underline{N}}} \Bigg)^2 \\
\le & \Bigg(\sqrt{P^{\mathcal{P}}_{ph,av,x}} \sqrt{1 - 2^{-\underline{N}}} \\
& \quad + \sqrt{1 - P^{\mathcal{P}}_{ph,av,x}} \sqrt{2^{-\underline{N}}} \Bigg)^2,
\end{aligned}$$

where the last inequality follows from the concavity of $\left(\sqrt{P_{ph}}\sqrt{1-2^{-t}} + \sqrt{1-P_{ph}}\sqrt{2^{-t}}\right)^2$ concerning P_{ph}.

[1] C.H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," Proc. IEEE Int. Conf. on Computers, Systems, and Signal Processing (Bangalore, India, IEEE, New York, 1984) 175.

[2] D. Mayers, "Quantum key distribution and string oblivious transfer in noisy channels," In *Advances in Cryptology – Proc. Crypto'96, Vol. 1109 of Lecture Notes in Computer Science* (Ed. N. Koblitz, Springer-Verlag, New York, 1996) 343; J. Assoc. Comput. Mach. 48 (2001) 351.

[3] W.-Y. Hwang, "Quantum Key Distribution with High Loss: Toward Global Secure Communication," *Phys. Rev. Lett.*, **91**, 057901 (2003).

[4] H.-K. Lo, X.-F. Ma, and K. Chen, "Decoy State Quantum Key Distribution," *Phys. Rev. Lett.*, **94**, 230504, (2005).

[5] X.-F. Ma, B. Qi, Y. Zhao, and H.-K. Lo, "Practical decoy state for quantum key distribution," *Phys. Rev.* **A 72**, 012326 (2005).

[6] X.-B. Wang, "Beating the PNS attack in practical quantum cryptography," *Phys. Rev. Lett.*, **94**, 230503 (2005).

[7] H. Inamori, N. Lütkenhaus, and D. Mayers "Unconditional Security of Practical Quantum Key Distribution," quant-ph/0107017.

[8] S. Watanabe, R. Matsumoto, and T. Uyematsu, "Noise Tolerance of the BB84 Protocol with Random Privacy Amplification," quant-ph/0412070.

[9] R. Renner, N. Gisin, and B. Kraus, "Information-theoretic security proof for quantum-key-distribution protocols," *Phys. Rev.* **A72** (2005) 012332, quant-ph/0502064.

[10] M. Hayashi, "Practical Evaluation of Security for Quantum Key Distribution," *Physical Review A*, **74**, 022307 (2006).

[11] P. W. Shor and J. Preskill, "Simple Proof of Security of the BB84 Quantum Key Distribution Protocol," *Phys. Rev. Lett.* **85**, 441 (2000).

[12] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, "Security of quantum key distribution with imperfect de-

vices," *Quant. Inf. Comput.*, **5**, 325 – 360 (2004).
[13] H.-K. Lo, "Getting something out of nothing," quant-ph/0503004.
[14] J.-C. Boileau, J. Batuwantudawe, and R. Laflamme "Higher-security thresholds for quantum key distribution by improved analysis of dark counts," *Phys. Rev.*, **A 72**, 032321 (2005).
[15] M. Hayashi, A. Tomita, T. Hiroshima, and J. Hasegawa, "Security analysis of quantum key distribution system based on decoy method incorporating finite statistics," in preparation.
[16] M. Hayashi, "General theory for decoy-state quantum key distribution with arbitrary number of intensities," quant-ph/0702251.
[17] L. Carter and M. Wegman. Universal classes of hash functions. Journal of Computer and System Sciences, Vol. 18, No. 2, 1979, pp. 143–154.
[18] H. Krawczyk. LFSR-based hashing and authentication. Advances in Cryptology — CRYPTO '94. Lecture Notes in Computer Science, vol. 839, Springer-Verlag, pp 129–139, 1994.
[19] M. Koashi, "Efficient quantum key distribution with practical sources and detectors," quant-ph/0609180.
[20] A. S. Holevo, "Covariant measurements and uncertainty relations," *Rep. Math. Phys.*, **16**, 385–400, (1979).
[21] There is a possibility that to improve this bound if our two-way error correction code is chosen carefully.